

FIGURE 1

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCTTCTGCATCGCGGCTTCGGCG
GCTTCCACCTAGACACCTAACAGTCGCGGAGCCGCCGCTCGTGAGGGGGTCGGCACGGGGAGTCGGGCGGTCT
TGTGCATCTTGGCTACCTGTGGGTCTGAAGATGTCGGACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGC
GCTATTGGTTCGCCGCCACCGTCGCCCTTGGTTCGGCAAACCTCGGCCTCATCAGCCCGGCCCTACCTCTTCC
TCTGGCCCGAAGCCTTCTTTATCGCTTTCAGATTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGGGT
CAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGACTTGAAACAGGAGCTT
TTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTAACTGGATTGTCATCGTGATTACTGGCTTAGCAA
TGGATATGCACTTGTCTGATGATTCCTCTGATCATGTGACTTTATGTCTGGGCCAGCTGAACAGAGACATGA
TTGTATCATTTTGGTTTGGAACACGATTTAAGGCCCTGCTATTTACCTGGGTTATCCTTGGATTCAACTATATCA
TCGGAGGCTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTCTTAATGTTTCAGATACC
CAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGGCTGCCCAGTAGGAGAGGAG
GAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGCTGCTGATCAGAATGGCGGAGGCGGGAGACACA
ACTGGGGCCAGGGCTTTCGACTTGGAGACCAGTGAAGGGGCGGCCCTCGGGCAGCCGCTCCTCTCAAGCCACATTT
CCTCCAGTGCTGGGTGCACTTAACAACTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTC
TTTCAGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCACAAGTTTCACGATTCTCATTCA
AGTCTTTACTGCTGTGAAGAACAAATACCAACTGTGCAAAATTGCAAAACTGACTACATTTTGGTGCTTCTCTCT
TCTCCCCCTTCCGCTCTGAATAATGGGTTTGTAGCGGGTCTTAATCTGCTGGCATTGAGCTGGGGCTGGGTCAACAA
ACCTTTCCTCAAAAGGACCTTATCTCTTTCTTGCACACATGCCTCTCTCCACTTTTCCCAACCCCCACATTTGCA
ACTAGAAAAAGTTGCCCATAAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAAGGCTGGT
CACAACAATCATATTCAGTTATTTTCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGGAGAAGACAGCCACG
GATGAAGCGTTTCTCAGCTTTTGGAAATGCTTCGACTGACATCCGTTGTTAACCGTTTGCCACTCTTCAGATATT
TTTTATAAAAAAAGTACCACCTGAGTTCTAGGGGCCACAGATTGTTTATTAATGAGATACGAGGGTTGGTGCTGG
GTGTTTGTTCCTGAGCTAAGTGATCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGG
GGATGCACCCCTTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGTTA
GGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTGATTTCCATTTCAATC
TCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCTCATACGCTATTTAAATGTCACTTTTTTGCCTA
TCCCCCGTTTTTGGTTCATGTTTCAATTAATTGTGAGGAAGGCGCAGCTCCTCTCTGACAGTAGATCATTTTTTA
AAGCTAATGTAAGCACATCTAAGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGGAG
GTGTGTTATTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTTTCGTA
GGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCATGGCTTTTCACACAGTTA
TTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGTATTGAGTGGCTGTCACTTTGAGGCAACTA
AAAAGGCTTCAAACGTTTGTATCAGTTTCTTTTTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCC
ACTCTTAAACAGTGTGATGTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTT
GTGTGTAATCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACCTGTCATGTTTTGT
TGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAAATTTAGAGGCCATTACTCCAATTATGTTGCACGTACAC
TCATTGTACAGGCGTGAGACTCATTGTATGTATAAGAATATTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGT
ACCTCTTACCAGTCAGCTGCCTGCGAGCAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAATTTTTCAGTT
CAGGGCAAAATGTTTCATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGA
TTATGTTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTGTGACTTGACAAGGCAAAACATGACAGTGGAT
TCTCTTTACAAATGGAATAAAAAAATCCTTATTTTGTATAAAGGACTTCCCTTTTTGTAAACTAATCCTTTTTAT
TGGTAAAAATGTAAATTAAATGTGCAACTTG

2/615

FIGURE 2

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYFPVGPGTGFLYLVN
LYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIPLIMSVLYVWAQLNRDMIVSFWFGTRF
KACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFLMFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPA
SMRRAADQNGGGGRHNWGQGFR LGDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

100556 0130
200610 09525001

FIGURE 3

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTCAAATGCACGGATT
CTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGAAGTCCTGAGCGCTCAAGTTTGTCCGT
AGGTCCGAGAGAAGGCC**ATGG**GAGGTGCCGCCACCGGCACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCC
CCGAGTCTTTGCTGCCGAAGCTGTGACTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGT
CCCAGAGCCCTATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAAT
TTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAATACC
AGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCATAACCGGTTTGATGCTGT
GCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTTCGTTATGGCTGGCGCTGGGGTTGGAGAAGTGCAGTGTT
TGTGACTATATTCAACACAGTGAACACTAGTCTGAATGTATAACGAAATAAAGATGCCTTAAGCCATTTTGTAAAT
TGCAGGAGCTGTCACGGGAAGTCTTTTATAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCATAATTGG
AGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTGGTGAGACTGTTTACAGGAAAG
AAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAAGAGTGGAAAGGCAGACTACAAGTTACTGAGCA
CCTCCCTGAGAAAATTGAAAGTAGTTTACGGGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCT
AAACCTTCCTAGAAACCTTTCAGTAATAGATAAAACAAGACAAGGAC**TGA**AAGTGCTCTGAACCTGAAACTCACTG
GAGAGCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGCTGACAAAATT
TAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTCTTTTCTTTTAACTAAGAATGGGGCTGTT
GTACTCTCACTTTACTTATCCTTAAATTTAAATACATACTTATGTTTGTATTAATCTATCAATATATGCATACAT
GGATATATCCACCCACCTAGATTTTAAAGCAGTAAATAAAACATTTTCGCAAAGATTAAAGTTGAATTTTACAGTTT

4/615

FIGURE 4

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELFGKDEQQRISKDLA
NICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSOAEIYHNRFDVQSAHRAATRGFIRYGWRWGWRTAVFVTIFN
TVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLRGLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDR
KALHELKLEEWKGRQLQVTEHLPEKIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 5

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGCGCCTGAAGTCGGC
GTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGCTTAAGAAGTAAAAATGGCAGGCTTCC
TAGATAATTTTCGTTGGCCAGAATGTGAATGTATTGACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTCTG
CAGGTATATTGTTTTTTTACAGGCTGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGA
ACCATGCCTTTACACATGTGGTGTATTTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTCAGG
TGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTTTTCATTGGTTTCATGT
TGATGTTTGGGTCACCTTATTGCTTCCATGTGGATTCTTTTTGGTGCATATGTTACCCAAAATACTGATGTTTATC
CGGGACTAGCTGTGTTTTTTTCAAATGCACTTATATTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAG
AGCTATGGACCCTGAGATCACTTCTTAAGTCACATTTTCCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATC
TCTCAGTACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTTATGTTCTGAGTT
TTGAAATAGTTTTATGAAATTTCTTTATTTTTTTCATTGCATAGACTGTTAATATGTATATAATACAAGACTATATG
AATTGGATAATGAGTATCAGTTTTTTTATTCTGAGATTTAGAACTTGATCTACTCCCTGAGCCAGGTTACATCA
TCTTGTCTATTTAGAAAGTAACCACTCTTGTCTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTT
TGGGAGGCCGAGGCGGGCCGATTGCTTGAGGTCAAGTGTGTTGAGACCAGCCTGGCCAACATGGCGAAACCCCATC
TACTAAAAATACAAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGCTGAGGCAGG
AGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCACTGCACTCTAGCCTGGGGGAGAA
AGTGAAACTCCCTCTCAAAAAAAGACCACTCTCAGTATCTCTGATTTCTGAAGATGTACAAAAAATATAGCTT
CATATATCTGGAATGAGCACTGAGCCATAAAAGGTTTTTCAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGT
TTTTTTGGTAAAGAAAAAATATTTGTTCTTATGTATTGAAGAAGTGACTTTTATATAATGATTTTTTAAATGCC
CAAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCTCTAATATGACTTTATGTGAGAA

6/615

FIGURE 6

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWIMIDAADVYPKPEQLNHAFHTCGVFSTLAFMINAV
SNAQVRGDSYESGCLGRTGARVWLFIFGMLMFGSLIASMWILFGAYVTQNTDVYPGLAVFFQNALIFFSTLIYKF
GRTEELWT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 23-42 (type II), 60-80, 97-117, 128-148

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CGCTGGTTTTTGTCTCTGCAATAGGCGGGCTTAGAGGGAGGGGCTTTTTTCGCCCTATACCTACTGTAGCTTCTCCACG
TATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAGGATGTGCGTTCTTC
CACTAGAGAAGCTCTTCTGAGGGAGGTAATTAAAAAACAGTGGGAATGCAAAAAACAGTGCTGTAGTCATCCTGTAATA
TGCTCCTTGTCAACAATGTATACATCTCTGCTAGGTGCCATATTCCATTGCTTTAAGCTCAAGTCGCATCTTACTA
GTGAAGTATTCTGCCAATGAAGAAAAACAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAAGCTGGTG
AAGCTAGTTTTCTGTGTGCTTGTGTCACTTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATATGCT
TCCTGGAAGGAATTCCTGATTTTCATGAAGTGGTCCATTCTGCTTTCTTTATTTCTGGATAACTTTGATTGTG
TTCTATGTCCTGTCCCTATCTTCAACCAGCCATGGCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTT
CTATTTCAGGATAGTGTCTGAAGAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTTATTTTGTCTATT
GTGGCCTTGACTGCCGGGACTAAAACTTTACAGCACAACCTTGGCAGGACGTGGATTTTCATCAGCATGCCCTTTCTC
AGCCCTTCCAATTCCTGCTTCTTTTTCAGAAGTGAGTGTCCAGAAAAAGACAATTTGTACAGCAAAGGAATGGACT
TTTCTTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTTCAGTCACATCCGCTCTTGGCATGGGCCATGTTCTTATT
ATAGTCCAGTGTTTTTATTTCTTCAATGGCTAATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACT
GAAAGCATCTTCTCATACAGAACAGCAAACCTCTATTTCTTTGGCATTCTGTTAATGGGCTGACTCTGGGCCTTCAG
AGGAGTAACCGTGATCAGATTGAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTTTCAGTAGCCCTTATTTTT
GTAAGTGCATTCCAGGGCCTTTTCAGTGGCTTTTCATTCTGAAGTTCCCTGGATAACATGTCCATGTCTTGATGGCC
CAGGTTTACCAGTGTCAATTATCACAACAGTGTCTGTCTGGTCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTG
GAAGCCCCCATCAGTCCCTTCTCTCTATATTTATTTTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCACCTAGG
CAAGAAAGGATCCGAGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAAGACTT
ACCAAACCCAGAGTGATGAGTCAGATGAAGATACTTTCTAACTGGTACCCACATAGTTTTGCAGCTCTCTTGAAC
CTTATTTTTACATTTTTTCAGTGTTTTGTAAATATTTATCTTTTTTCATTTTGATAAAACCAGAAATGTTTTCTAAATCCTAA
TATTTCTTTGCATATATCTAGCTACTCCCTAAATGGTTCCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTC
TAAAGAACTGATACAGGAGTAACAATATGAAGAAATTCATTAATATCTCAGTACTTGTATAAATCAGAAAGTTATAT
GTGCAGATTATTTTTCTTGGCCTTCAAGCTTCAAAAAAACTTGTAAATAATCATGTTAGCTATAGCTTGTATATAC
ACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGTCTTCCCTTTTTTAACA
TTATAAAAAGCTAGGTTGTCTCTTGAATTTTTTGAGGCCCTAGAGATAGTCAATTTTGCAAGTAAAAGCAACGGGACC
CTTTCTTAAAAACGTTGGTTGAAGGACCTTAAATACCTGGCCATACCATAGATTTTGGGATGATGAGTCTGTGCTAA
ATATTTTGTCTGAAGAAGCAGTTTTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAATTTGG
ATTTTTTGTAAATAATCTTTTGATGTTTTTAAACATTGGTTCCCTAGTCACCATAGTTTACCACCTTGTATTTTTAAGTCA
TTTAAACAAGCCACGTTGGGCTTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGATGTCATTACTCCTGAATTA
TTACATTTTGGAGAATAAGAGGGCATTTTATATTTTATTAGTTACTAATTCAAGCTGTGACTATTGTATATCTTTCC
AAGAGTTTGAATGCTGGCTTCAGAATCATACCATAGTTGTCACTGAAGCTGATGCCTAGGAACCTTTTAAAGGGATC
CTTTCAAAGGATCACTTAGCAAAACACATGTTGACTTTTTAAGTCACTGATGTATGAATTAATACTCTAAAAATAGAA
AGACCAGTAATATATAAGTCACTTTACAGTGCTACTTTCACACTTAAAAGTGCATGGTATTTTTTCATGGTATTTTG
CATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAATAAATTAGCAAAACAAAAGTGAC
TTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGGCTTTAACTGGCAGGCCTGTATGTTTACAGACTAC
CATACTGTAAATATGAGCTTTATGTTGTCTATTCTCAGAAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTC
ATGCAGATGAATATAAGGTAATATACATTTATATAAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAAT
TGGTGGAAATTTGTAAATTAATAAATATTTAAACCT

FIGURE 8

MEKQCCSHPVICSLSTMYTFLLGAI FIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVKLVFCVLVSFCVIK
KDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQ PAMAVIFS NFSIITTALLFRIVLKRRLNWIQ
WASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFFSPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFS
HIRLGMGHVLIIVQCFISSMANIYNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFY
GHSAFSVALIFVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFIYNA
SKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318,
328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

GGGGCTTCGCGGCCAGCGGCCAGCGCTAGTTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATGAGCAGGTCTG
AAGACTAACATTTTGTGAAGTTGTAAAAACAGAAAACCTGTTAGAAATGTTGGTGGTTTTAGCAAGGCCCTCAGTTTTC
CTTCTCTCAGCCCTTGTAAATTTGGACATCTGCTGCTTTTCATATTTTTCATACATTACTGCAGTAACACTCCACCA
ATAGACCCGGCTTTTACCTTTATATCAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGCAATGAGTA
AATATTGCGGCAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTTCATGCTCTGAGTCTGATGCTGAAGAG
AACGTTTATCATCAAATTAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCA
AACTTCCAGAAAACAACCTTTTGTCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCATTATAT
ATGTTTTGTTTCAGACCATCTCTTCTCAACAAATGCAAGCCCAAAATCCATGGCAAAACAAGTCTTCTGGATCAGACT
TTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATCTGACTGCTATGCTTTTCAGACTGGCAATTTT
GGGACTGATTTTAGAACAGAAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTTCATGATCACTACTGCA
GCAGAAATGGTCTATGTCATTTTCTTCTTTGGTTTTTTCTTGACTTACATTTCGTGATTTTTCAGAAAATTTCTTTA
CGGGTGAAGCCAATTTACATGGATTAAACCTCTATGACACTGCACCTTGCCCTATTAAACAATGAACGAACACGG
CTACTTTCCAGAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTACACAGAAGTTGCTTATTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCT
GATAATCAGGAACATGAAAGAGCCATTTGTATGATTATTCTAAAGGATATCATCAAGAAGACTATTTAAAAACA
CCTATGCCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACTATG

10/615

FIGURE 10

MWWFQQGLSFLPSALVIWTSAAFI FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAVLCIATIVVRY
KQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFGMGSLYMFVQTILSYQMOPKI
HGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT
YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTRLLSRDI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

CCACACGCGTCCGCCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCGCGGGGCCGGGGTGCGGAG
CCGACATGCGCCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCACCTTCGCCTTGTACTTGTCTGTGACGCG
GACTGCCCGCGCGGGCGGAGACTGGGCTCCACCGAGGAGGCTGGAGGCAGGTGCTGTGGTTCCCCTCCGACCTGG
CAGAGCTGCGGGAGCTCTCTGAGGTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCCTACGTGTTCTTGCTCTTCT
CGCGCGCTACCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGGTTCTCTGAATGTTTTAGCTGGTGCCCTTGT
TTGGGCCATGGCTGGGGCTTCTGCTGTGTGTGTGTGACCTCGGTGGGTGCCACATGCTGCTACCTGCTCTCCA
GTATTTTTTGGCAAACAGTTGGTGGTGTCTACTTTTCTGATAAAGTGCCCTGCTGCAGCAAAAGGTGGAGGAGA
ACAGAAACAGCTTGTTTTTTTTCTTATTGTTTTTTGAGACTTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCT
CGGCCCAATTCTGAACATTCCCATCGTGCACTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCA
TCTGTGTGCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTGTCTTTA
AGCTGTGTGGCCATTGCCATGTGTGCATTAATTCTTGAACCCCTCATTAAAAAATTTAGTCAGAAACATCTGCAAT
TGAATGAAACAAGTACTGCTAATCATATACAGTAGTAAAGAAGACATGTACTGGATTTTCTGTTTGGCCACTCC
CTGGACTCAGTTGCTTATTTGTGTAATGGATGTGGTCCCTCTAAAGCCCCCTATTGTTTTTGTATTGTCCTCTATAG
GTGATGTGGACACTGTGCATCAATGTGCAGTGTCTTTTCAGAAAAGGACACTCTGCTCTTGAAGGTGTATTACATC
AGGTTTTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCCTCTAGAAAAATGCTGTTTGTGGCCGGGCG
CGGTGGCTCACGCCGTGAATCCAGCACTTTGGGAGGCCGAGGCCGGTGATTACAAAGGTGAGGAGTTCAAGACC
AGCCTGGCCAAGATGGTGAATCCTGTCTCTAATAAAAAATACAAAAATAGCCAGGCGTGGTGGCAGGCACCTGT
AATCCCACTACTCGGAGGCTGAGGCAGGAGAATTGCTTGAAACCAAGTGCCAGAGGTTGCAGTAAGCCAAGAT
CACACCAGTGCATCCAGCCTGGGTGATAGAGTGAGACCTGTGCTTGAC

12/615

FIGURE 12

MRPLLGLLLVFAGCTFALYLLSTRLLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYRKEHQAYVFLLFCG
AYLYKQGFAIPGSSFLNVLGALFGPWLGLLLCCVLTSVGATCCYLLSSIFGKQLVVSYPDKVALLQRKVEENR
NSLFFFLFLRLFPMTNPNWFLNLSAPILNIPVQFFSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKL
LAIAMVALIPGTLLIKKFSQKHLQLNETSTANHIHSRKDT

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 13

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACCGTGTGAG
GGGGCCTGTGGCCCCAGCGTGTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTCCTTACACTTCGCCA
TGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTGGATTGGGGTGGCTTTTCTTCA
TGCGCCAATTGTTTAAAGACTATGAGATACGTGAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTT
CTTGCACCATGTTTGGAGTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGA
AAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTACATTGGCTATTTTATGTGAGCA
ATATCCGACTACTGCATAAAACAACGACTGCTTTTTTCCGTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGA
AACTAGGAGATCCCTTTCCCATTTCTCAGCCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTG
GTGTGATTGGAGTGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCCATACACTTACATGTCTT
ACTTCCTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGATCA
TAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACAAACCATCAG
GTTTCTGGGGAAATGATAAAAAGTGTACCCTTCAGCATCAGGAAGTGAAAATCTTACTCTTATTCAACAGGAAG
TGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAG
AATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGAAAA
TTTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGCATTGAGATCA
CTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATTTCTTTCATTCTTGTGGAATAA
TCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCT
CCAATGTCATTGTCTGCTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGA
GTATGCCTTTAGAATACCGCACCATTAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGT
TTGATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCCAGAGA
AGCAAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTAGATATAAGAGG
GGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAACAAACAAAATGCTATGGTAGCATTTTTTCACCTTCATAGC
ATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAA
GACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGCCAAGAACTAAAG
GTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTA
AGGTTCCATGGAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAATCAGAGACTGTAACAAAAA
AAAAAAAAAAGGGCGCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCATGGCCCACTTGTATTATG
CAGCTTATAATG

14/615

FIGURE 14

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFASFCTMFELIIFEILGVLNSSSRYPFW
KMNL CVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDPPILSPKHGILSIEQLISRV
GVIGVTL MALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPS
GFWGMIKSVTTSASGSENLT LIQQEVD ALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWK
IFMATINIVFDRVGKTD PVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKEFFYAISSKS
SNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPE
KQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, 425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 15

GACGGAAGAACAGCGCTCCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCGGGACATGCGGCCC
CAGGAGCTCCCCAGGCTCGCGTTTCCCGTTGCTGCTGTTGCTGTTGCTGCTGCTGCCGCCGCCGCGTGCCTTGCC
CACAGCGCCACGCGCTTCGACCCACCTGGGAGTCCCTGGACGCCCCCAGCTGCCCGCGTGGTTTTGACCAGGCC
AAGTTCGGCATCTTCATCCACTGGGGAGTGTTCCTGCGCCAGCTTCGGTAGCGAGTGGTCTGGTGGTATTGG
CAAAAGGAAAAGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGATTTT
GGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTTCAGGCCCTCGGTGCCAAATAC
ATTGTCTTAACCTCCAAACATCATGAAGGCTTTACCTTGTGGGGGTGAGAATATTTCGTGGAACCTGGAATGCCATA
GATGAGGGGGCCCAAGAGGGACATTGTCAAGGAACTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGA
CTGTACTATTTCCCTTTTTGAATGGTTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAA
TTTCCAGTTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATATCAGCCTGAGGTTCTGTGGTCCGAT
GGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTTATATAATGAAAGCCAGTT
CGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCATCTGTAAGCATGGTGGCTTCTATACCTGCAGT
GATCGTTATAACCCAGGACATCTTTTGCCACATAAATGGGAAAACCTGCATGACAATAGACAACTGTCTGGGGC
TATAGGAGGGGAAGCTGGAACTCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCA
TGTGGAGGAAATCTTTTGATGAATATTTGGGCCACACTAGATGGCACCATTCTGTAGTTTTTTGAGGAGCGACTG
AGGCAAGTGGGGTCTTGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTATACCTGGCGATCCAGAATGAC
ACTGTCACCCAGATGTGTGGTACACATCCAAGCCTAAAGAAAAATTAGTCTATGCCATTTTTCTTAAATGGCCC
ACATCAGGACAGCTGTTCTTGGCCATCCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGA
CAGCCACTTAACTGGATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTTCATCAGATG
CCGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTGATGCTGCAAGTT
ATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGAAAAGCAATGTAAACTGGATAAGAAAA
TTATTTGGCAGTTTCAGCCCTTTCCCTTTTTCCCACTAAATTTTTCTTAAATTACCCATGTAACCATTTTTAACTCT
CCAGTGCACCTTTGCCATTAAAGTCTCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCA
CATTATAGTAGCAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTTATGTTGAAGCCATATCCCCCATGAT
TATATAGTTATGCATCACTTAATATGGGGATATTTCTGGGAAATGCATTGCTAGTCAATTTTTTTTTTGTGCCAA
CATCATAGAGTGTATTTACAAAATCCTAGATGGGCATAGCCTACTACACACCTAATGTGTATGGTATAGACTGTTG
CTCCTAGGCTACAGACATATACAGCATGTTACTGAATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAA
ACATATGGAAACATAGAGAAGGTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCAC
GAATGGAGCTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTATTGAA
CACTGCCAGACGTTATAAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAGTTTTTCTTTCTTCAATT
ATAAATTAAACATAAGTGTACTGTAACCTTACAAACGTTTTTAATTTTTTAAACCTTTTTGGCTCTTTTGTAATAAC
ACTTAGCTTAAACATAAACTCATTGTGCAATGTAA

FIGURE 16

MRPQELPRLAFPLLLLLLLLLLLLLPPPPCPAHSATRFDPDTWESLDARQLPAWFDQAKFGIFIHGWFVSVPSPFGSEWFW
WYWQKEKIPKYVEFMKDNYPPSPFKYEDFGPLFTAKFFNANQWADIFQASGAKYIVLTSKHHEGFTLWGSEYSWNW
NAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEWFHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVL
WSDGDGGAPDQYWNSTGFLAWLYNESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKL
SWGIRREAGISDYLTIELVKQLVETVSCGGNLLMNI GPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYTWRS
QNDTVTPDVWYTSKPKEKLVYAI FLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI SLEQNGIMVELPQLTI
HQMPCKGWGALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

[illegible]

18/615

FIGURE 18

MKFLLDILLLLLPLLIIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSKL
VLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAIEIGDVSILVNNAGVVYT
SDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYCSSK
FAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMHGILT
EQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

20250310 09:25:00

19/615

FIGURE 19

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGAT
GCTGCTGCCGCGGTTGCAGTTGTCTGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGTAG
CGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGGCCATGAGGCCGGAACCGCGCTACAGGCCG
TGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTCTGCCTGCTGAGTGCCTCGG
ATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTATAAAG
TCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCTGCAGGA
GGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAAAAGTTCA
TTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGAGGAGAAAC
AAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGAATGATGGCAGCATATCACAAATTTA
GGAATGGTATGTGGATGAGCCGTCTGCGGCAGCGAGGTCTGCGTGGTTCATGTACCATCAGC
CATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGACCGGTGCAACA
TGAAGAACAATTTCAATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTTCTAGAGAAGCTG
AAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAGGAAGAAGATGCCA
AAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACATCCTAATCCCCAGCA
TTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTTGGATCTGTAGAAAAA
GAAAACGGGAGCAGCCAGACCTAGCACAAAGAAGCAACACACCATCTGGCCCTCTCCTCACC
AGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACAAAGCGAAGCTGACTTAG
CTGAGACCCGGCCAGACCTGAAGAATATTTTCATTCAGAGTGTGTTTCGGGAGAAGCCACTCCCG
ATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCATCAGAAAGTGGGTTTGTGACTC
TGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGAGTTCTCCCCAGACCAAATGG
GGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATGGTTATTAGGACATATAAAAAA
CTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATCCTCTTATTTTCTATAAGGAAAA
TACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGTGGATGAGCATGTGGTCCCCACGA
CCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTATCCCAGCCAGTCATCCAGCTCGACC
TTATGAGAAGGTACCTTGCCAGGTCTGGCACATAGTAGAGTCTCAATAAATGTCACTTGTT
GGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTGGCAGTGATAAAGATGGGCTGTGGAGC
TTGGAAAACACCTCTGTTTTCTTGTCTATACAGCAGCACATATTATCATAACAGACAGAAA
ATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAGGTTGGCCTGTGCATCGGCAATTCTC
ATATCTGTTTTTTTTCAAAGAATAAAATCAAATAAAGAGCAGGAAAAAAAAA

205356.01501

FIGURE 20

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRRGGTQRPCYKVIYFHDTSRRLN
FEEAKEACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLA
WTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE
KPAVPSREAEGEETELTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLLVTTV
VCWVWICRKRKREQPDSTKKQHTIWPSPHQGNPDLEVYNVIRKQSEADLAETRPDLKNISF
RVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENE
IYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

AGGCTCCCCGCGCGCGGCTGAGTGTGCGGACTGGAGTGGGAACCCGGGTCCCCGCGCTTAGAGAACACGCGATGACCA
CGTGGAGCCTCCGGCGGAGGCCGGCCCCGCACGCTGGGACTCCTGCTGCTGGTCTGCTTTGGGCTTCTTGGTGTCTCC
GCAGGCTGGACTGGAGCACCTTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA
ACTTTCATGCTGGAGGATTCCACCTTTCTGGATCTTTCGGGGGACTCCATCCACTATTTCGCTGTGCCAGGGAGTACT
GGAGGGACCGCTGCTGAAGATGAAGCCCTGTGGCTTGAACACCTTACCACCTATGCTTTCGCTGGAACTGCATG
AGCCAGAAAGAGGCAAAATTGACTTCTCTGGAACTGGACCTGGAGGCCCTTCTGCTGATGGCCCGAGAGATCG
GGCTGTGGGTGATTCTGCGTCCAGGCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGGCCAGCTGGCTAC
TCCAAGACCTTGGCATGAGGCTGAGGACAACTTACAAGGGCTTACCAGAAGCAGTGGACCTTTATTTTGACCAC
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG
GTTCTCTATAATAAAGACCCCGCATACATGCCCTACGTCAGAAGGCATGGAGGACCGTGGCATTTGTGGAACCTGC
TCTTACTTTAGACAAACAGGATGGGCTGAGCAAGGGGATGTCCAGGGAGTCTTGGCCACCATCAACTTGCATG
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGATGGAGT
ACTGGACGGGGTGGTTTGAAGCTGCTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTTGAAAACCGTGT
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCCAACCTTTGGCTTCATGAATG
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG
ATTACACGGCCCAAGTACATGAAGCTTTCGAGACTTCTTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACTGTCTCTGTGGGACGCCCTCAAGTACC
TGGGGGAGCCCAATCAAGTCTGAAAAGCCCATCAACATGAGGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCCT
TCGGGTACATTTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG
TGTTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT
ACACCGTGTCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG
GCTTAATTGGAATCTCTATCTGAATGATTACCCCTGAAAAAAGTTCAGAATCTATAGCTTGGATTAAGAAGA
GCTTCTTTTCAAGAGTTTCGGCTCGGACCAATGNGNTTCCCTCCGAAAACCCACATTAACCTTCTTCTTGG
GTAGCTTTGTCCAGTCAAGTCCACGCTTGTGACAACTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGTTGTATTCA
TCAATGGCCAGAACCCTTGGACGTTACTGGAACATTGGACCCCCAGAAGACGCTTTACCTCCAGGTCCCTGGTTGA
GCAGCGGAATCAACCAGGTCTCGTTTTTGGAGAGACGATGGCGGGGCCCTGCATTACAGTTTACGGAAACCCCC
ACCTGGGCAGGAACCAGTACATTAAGTGAAGCGGTGGCACCCCTCCTGCTGGTGGCAGTGGGAGACTGCCGCCCTC
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG
ACTGGGGCTTACAGTCTGCCCTGTCTCAGCTCAAAACCTTAAGCTTGCAGGAAAGGTGGGATGGCTCTGGGCC
TGGCTTTGTTGATGATGGCTTTCTTACAGCCCTGCTCTTGTGCGAGGCTGTCTGGGCTGTCTTAGGTTGGGAGC
AGCTAATCAGATCGCCCAGCCTTTGGCCCTCAGAAAAAGTGC'TGAAACGTGCCCTTGCACCGGACGCTCACAGCCC
TGCGAGCATCTGCTGGACTCAGGCGTGTCTTTTGTCTGGTTTCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT
TTTATCCCCGAAATCCTGGGTGTGTACCAGTGTAGAGGGTGGGGAAGGGGTGTCTACCTGAGCTGACTTTGTT
CTTCTCTTCAACAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACTCGGCGTGAGAAACATGTGACTTCCCCCT
TCCCTTCCCACCTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCTTGGCAGAGCAAGCATGGCCCATGTCTGCA
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCC
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGGA
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG
ACAGAAGGCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCCCGAACAGCAGGGGGCAGAGCAGCCCTCCTTC
GAAGTGTGTTCCAAGTCCGATTGAGCCTTGTCTGGGGCCCAGCCCAACACCTGGCTTGGGCTCACTGTCTGTA
GTTGTCAGTAAAGCTATAAAGCTTGAATCACAA

FIGURE 22

MTTWSLRRRPPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFWI
FGGSIHYFRVPREYWRDRLKMKACGLNTLTITYVPWNLHEPERGKFDFSGNLDLEAFVLMAAE
IGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQYKR
GGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKGDLGSKGIVQGVLATINLQ
STHELQLLTTFLFNVQGTQPKMVMEYWTGWFD SWGGPHNILDSSSEVLKTVSAIVDAGSSINLY
MFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDY TAKYMKLRDFFGSISGIPLPPPD
LLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYETSITSSG
ILSGHVHDRGQVFVNTVVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGENIDDQRKGL
IGNLYLNDSP LKNFRIYS LDMKKSFFQRFGLDKWXS L PETPTLP AFFLGSLSISSTPCDTFLK
LEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPALQFTETPHLG
RNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 23

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAGA
AAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAA**ATGA**ATTCATCTAAATCATCT
GAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCTGGG
ATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCATCTTT
CAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTTACAGAGCTCTCCTGCTACAAT
TATGGATCAGGTTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCCAGCTGC
TACTTCTTTTCTACTGACACCATTTTCTGGGCGTTAAGTTTAAAGAACTGCTCAGCCATGGGG
GCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGAAACCTAAA
ATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTTGTGCGAGGGTCAGTGGCAATGGGTGGAC
GGCACACCTTTGACAAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAACATAGCTACC
CTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGAATGATGTAACC
TGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTGAACAAAGGAAAA
TCTCTT**TAAGA**ACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCAAGAACATGGCCAC
ACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTCATAAGTATTTGTTA
CTCTGATACAAATAAAAAATAAGTAGTTTAAATGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

24/615

FIGURE 24

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPENF
TELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQEF
LSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCAITMRDSSNPR
QNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:
amino acids 1-42

N-glycosylation site.
amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.
amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.
amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 27-38

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACCC
 TGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGGGG
 AGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATCCGA
 CACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCC**ATG**GCTCCCAAGA
 AGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCAGGCAG
 ACACTCGGTTCGTTTCGTAGTGGATAGGGGTTCATGACCGGTTTTCTCCTAGACGGGGCCCCGTTCC
 GCTATGTGTCTGGCAGCCTGCACTACTTTTCGGGTACCGCGGGTGCTTTGGGCCGACCGGCTTT
 TGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACCTACCACGAGC
 CACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGAATGAGGCAG
 CTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAGTGGGAGATGG
 GGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTTCATCTAAGAACCTCAGATCCAGACT
 TCCTTGCCCGCAGTGGACTCCTGGTTCAAGGTCCTTGCTGCCCAAGATATATCCATGGCTTTATC
 ACAATGGGGGCAACATCATTAGCATTTCAGGTGGAGAATGAATATGGTAGCTACAGAGCCTGTG
 ACTTCAGCTACATGAGGCACTTGCGCTGGGCTCTTCCGTGCACTGCTAGGAGAAAAGATCTTGC
 TCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGGGACTCTATACCACTG
 TAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTTCGGAAGTATGAACCCC
 ATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTACTGGGGCCAGAATCACT
 CCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGCTCAAGTTGGGAGCCAGTG
 TGAACATGTACATGTTCCATGGAGGTACCAACTTTTGATATTGGAATGGTGCCGATAAGAAGG
 GACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACCTATATCTGAAGCAGGGGACC
 CCACACCTAAGCTTTTTTGCTCTTCGAGATGTCATCAGCAAGTTCAGGAAGTTCCTTTGGGAC
 CTTTACCTCCCCCGAGCCCCAAGATGATGCTTGAGCCTGTGACTCTGCACCTGGTTGGGCATT
 TACTGGCTTTCTAGACTTGCTTTGCCCCCGTGGGCCCATTCATTCAATCTTGCCAATGACCT
 TTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACCGAACCTATATGACCCATACCATTT
 TTGAGCCAAACACCATTCCTGGGTGCCAAATAATGGAGTCCATGACCGTGCCCTATGTGATGGTGG
 ATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAGAGACAACTATTTTTTGACGGGGAAAC
 TGGGGTCCAAACTGGATATCTTGGTGAGAACATGGGGAGGCTCAGCTTTGGGTCTAACAGCA
 GTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTGGGGCAAACAATCCTTACCCAGTGGATGA
 TGTTCCCTCTGAAAATTGATAACCTTGTGAAGTGGTGGTTTTCCCTCCAGTTGCCAAAATGGC
 CATATCCTCAAGCTCCTTCTGGCCCCACATTCTACTCCAAAACATTTCCAATTTTAGGCTCAG
 TTGGGGACACATTTCTATATCTAACCTGGATGGACCAAGGGCCAAGTCTGGATCAATGGGTTTA
 ACTTTGGGCGGTACTGGACAAAGCAGGGGCCACAACAGACCCCTACGTGCCAAGATTCTCGC
 TGTTTTCTAGGGGAGCCCTCAACAAAATTACATTGACTGGAACTAGAAGATGTACCTCTCCAGC
 CCCAAGTCCAATTTTTTGATAAGCCTATCCTCAATAGCACTAGTACTTTGCACAGGACACATA
 TCAATTCCTTTTTCAGCTGATACACTGAGTGCCTCTGAACCAATGGAGTTAAGTGGGCAC**TGAA**
 AGGTAGGCCGGGCATGGTGGCTCATGCCTGTAATCCAGCACTTTGGGAGGCTGAGACGGGTG
 GATTACCTGAGGTGAGGACTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTA
 AAAATACAAAAATTAGCCGGGCGTGATGGTGGGCACCTCTAATCCAGCTACTTGGGAGGCTG
 AGGGCAGGAGAATTGCTTGAATCCAGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTG
 CACTCCAGCCTGGCTGACAGTGAGACACTCCATCTCAAAAAAAAAAAAAA

FIGURE 26

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRGHRFLLDGAPFRYVSGSLHYFRVPRVLW
ADRLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYICA
EWEMGGLPSWLLRKPEIHLRTSDPDFLA AVDSWFKVLLPKIYPWLYHNGGNIISIQVENEYGS
YRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDGFPADNMTKI FTLLR
KYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNM YMFHGGTNFGYWNG
ADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMMMLGPVTLH
LVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFM L YRTYMTHTIFEPTPFWVPNNGVHDRA
YVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNS SDFKGLLKPPILGQ TIL
TQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFY SKTFPILGSVGD TFLYLPGWTKGQVW
INGFN LGRYWTKQGPQOTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLDKPILNSTSTL
HRTHINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648**Tyrosine kinase phosphorylation site.**

amino acids 191-198

N-myristoylation site.amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 27

[illegible]

28/615

FIGURE 28

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGKDLLNGVKLVVETPEETLFTYQG
ASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQDKE
HDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQVCA
EQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPQPCGGPGLAPGVRSYGPRHRRL
HRYDVFCFATALKGRVYYLEHPEKLTLTAREACQEDDATIAKVGQLFAAWKFHGLDRCDAGW
LADGSVRYPPVHHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303,
307-313

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 29

GCAAGCGGCGAAATGGCGCCCTCCGGGAGTCTTGCAAGTCCCTGGCAGTCCTGGTGCTGTTG
 CTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTGCGGTCATCACGGACGAGAAC
 TGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCTTGT
 CAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAATATT
 GCGAAAGTAGATGTACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTCTTCCT
 ACTATTTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAGAAGGAC
 TTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATGGTTTGGT
 CCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGATCAGGACG
 TGCCATAACTACTTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACTGTTTTTGCT
 TTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGCAGATTGCCTT
 TGTCCTTCAAAAAGGCGCAGACCACAGCCATAACCATAACCCTTCAAAAAAATTATTATCAGAA
 TCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAAGATGTTTCAGAA
 GAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGCCATAAGACAACGC
 TCTCTGGGTCCATCATTGGCCACAGATAAATCC**TAG**TTAAATTTTATAGTTATCTTAATATTA
 TGATTTTGATAAAAACAGAAGATTGATCATTGTTTGGTTTGAAGTGAAGTGTGACTTTTTT
 GAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTACATTGAGAACATAAAA
 GCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATGGTTTAAATAGTTCTCTA
 ATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGTTTAAATAATAACCTATTT
 AAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTATTGAGGTATTTAAGAAGATT
 ATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTGTTTCACTGTGTGAAAAAAG
 AAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAGAAATGTGTATTTGAGTGACAA
 TTTTCGTGGTCTTTTTAGAGGTATATCCAAAATTTCTTGTATTTTAGGTTATGCAACTAAT
 AAAAATACTTACATTAATTAATTACAGTTTTCTACACATGGTAATACAGGATATGCTACTG
 ATTTAGGAAGTTTTTAAGTTTATGGTATTCTCTTGATTCCAACAAAGTTTGATTTTCTCTTGT
 ATTTTTCTTACTTACTATGGGTACATTTTTTATTTTTCAAATTGGATGATAATTTCTTGGA
 ACATTTTTTATGTTTTAGTAAACAGTATTTTTTGTGTTTCAAAGTTTACTGAGAGA
 TCCATCAAATTGAACAATCTGTTGTAATTTAAATTTTGGCCACTTTTTTCAGATTTTACATC
 ATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTTTTCTTTTTGGATGTGAAGGTGAACATT
 CCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTATTTTACATTTTGAAAATTCAAAGAAGCTT
 AATATAAAAGTTTGCATTCTACTCAGGAAAAAGCATCTTCTTGTATATGTCTTAAATGTATTT
 TTGTCCTCATATACAGAAAGTTCTTAATTGATTTTACAGTCTGTAATGCTTGATGTTTTAAA
 TAATAACATTTTTATATTTTTTAAAGACAACTTCATATTATCCTGTGTTCTTTCCTGACTG
 GTAATATTGTGTGGGATTTACAGGTAAAAGTCAGTAGGATGGAACATTTTAGTGTATTTTA
 CTCCTTAAAGAGCTAGAATACATAGTTTTTACCTTAAAGAAGGGGGAAATCATAAATACAA
 TGAATCAACTGACCATTACGTAGTAGACAATTTCTGTAATGTCCCCTTCTTCTAGGCTCTGT
 TGCTGTGTGAATCCATTAGATTTACAGTATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCC
 TTTAGAATTTAAATATTTGTACCATTAAAGAGTTTGGATGTGTAAGTTGTGATGCCTTAGAAA
 AATATCCTAAGCACAAAATAAACCTTTCTAACCACTTCATTAAAGCTGAAAAAAAAAAAAAA
 AAA

30/615

FIGURE 30

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNLQ
PEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFINF
ISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFALATL
FSGLLLGLCMIFVADCLCPSKRRRPQYPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSEEEAE
SKEGTNKDFPQNAIRQPSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 31

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCCGGTGCCGCGACTTTCACGATGGC
TCGCCCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAAAC
TGCCGCCGCTCTGCCACGGTCTGCCACCCCAACGCGAAGACGGTAACCCGTGTGACTTTGACT
GGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGATCCA
TCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAATTCTTT
TCTTCCGCTTGGATATTTCGCATGGGCCCTACTTTACATCACACTCTGCATAGTGTTTCTGATGA
CGTGCAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTG
ATGAGGAAC TAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAATTGGTCTA
ATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAACGTGTACAGGGC
TAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTACAAAGTGAGCA
CATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAAGGAGGCAATGC
GGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGA
TCCGAGAATTTAACTTAAATGAGCTATAACCAGCGGGCCAAGAACTATCAAAGGCTGGAGACA
ATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGATGGGGAAAACAAGA
AGGATAAAATAAAGATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATTCAGGCTCTTTCCAT
AACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGGCTGNGACTGGNTGGGG
CAGCATGCAGCTTCTGATTTTTAAAGAGGCATCTAGGGAATTGTCAGGCACCCTACAGGAAGGC
CTGCCATGCTGTGGCCAACGTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGAGGGGGA
AATGGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATCAGCTATTGAGACATCTCCA
TGGTTTCTCCATGAAACTCTGTGTTTTCATCATTCCTTCTTAGTTGACCTGCACAGCTTGGTT
AGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAACGCTAAGAATTTTCCCCAAG
GACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTCTTCATTAAAAGTATAAGCCTAA
CTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAGTTTTTATCATTGAAGACAATATT
GAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAATAGAGGCTTGAGACTTTCCTTTGT
GTGGTAGGACTTGGAGGAGAAATCCCTGGACTTTCACTAACCCTCTGACATACTCCCCACAC
CCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATTTTCCTTTTG

FIGURE 32

MAVLAPLIAIVYSPRLSRWLAQPYLLSALLSAAFLLVRLPPLCHGLPTQREDGNPCDFDW
REVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFLMT
CKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPYADLSLKYNCTGL
NFGKVDVGRYTDVSTRYKVSTSPILTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSEENVI
REFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

FIGURE 33

CGGACGCGTGGGGTGCCCCGAC**ATG**GCGAGTGTAGTGCTGCCGAGCGGATCCCAGTGTGCGGCG
GCAGCGGCGGCGGCGGCGCCTCCCGGGCTCCGGCTTCTGCTGTTGCTCTTCTCCGCCGCGGCA
CTGATCCCCACAGGTGATGGGCAGAATCTGTTTACGAAAGACGTGACAGTGATCGAGGGAGAG
GTTGCGACCATCAGTTGCCAAGTCAATAAGAGTGACGACTCTGTGATTTCAGCTACTGAATCCC
AACAGGCAGACCATTTATTTTCAGGGACTTCAGGCCTTTGAAGGACAGCAGGTTTTCAGTTGCTG
AATTTTTCTAGCAGTGAACCTCAAAGTATCATTGACAAACGTCTCAATTTCTGATGAAGGAAGA
TACTTTTGCCAGCTCTATACCGATCCCCCACAGGAAAGTTACACCACCATCACAGTCCTGGTC
CCACCACGTAATCTGATGATCGATATCCAGAAAGACACTGCGGTGGAAGGTGAGGAGATTGAA
GTCAACTGCACTGCTATGGCCAGCAAGCCAGCCACGACTATCAGGTGGTTCAAAGGGAACACA
GAGCTAAAAGGCAAATCGGAGGTGGAAGAGTGGTCAGACATGTACACTGTGACCAGTCAGCTG
ATGCTGAAGGTGCACAAGGAGGACGATGGGGTCCAGTGATCTGCCAGGTGGAGCACCTGCG
GTCACTGGAAACCTGCAGACCCAGCGGTATCTAGAAGTACAGTATAAGCCTCAAGTGCACATT
CAGATGACTTATCCTCTACAAGGCTTAACCCGGGAAGGGGACGCGCTTGAGTTAACATGTGAA
GCCATCGGGAAGCCCCAGCCTGTGATGGTAACTTGGGTGAGAGTCGATGATGAAATGCCTCAA
CACGCCGTACTGTCTGGGCCCAACCTGTTTCATCAATAACCTAAACAAAACAGATAATGGTACA
TACCGCTGTGAAGCTTCAAACATAGTGGGGAAAGCTCACTCGGATTATATGCTGTATGTATAC
GATCCCCCACAACCTATCCCTCCTCCCACAACAACCACCACCACCACCACCACCACCACC
ACCATCCTTACCATCATCACAGATTCCCGAGCAGGTGAAGAAGGCTCGATCAGGGCAGTGGAT
CATGCCGTGATCGGTGGCGTCGTGGCGGTGGTGGTGTTCGCCATGCTGTGCTTGCTCATCATT
CTGGGGCGCTATTTTGCCAGACATAAAGGTACATACTTCACTCATGAAGCCAAAGGAGCCGAT
GACGCAGCAGACGCAGACACAGCTATAATCAATGCAGAAGGAGGACAGAACAACCTCCGAAGAA
AAGAAAGAGTACTTCATC**TAG**ATCAGCCTTTTTGTTTCAATGAGGTGTCCAACCTGGCCCTATT
TAGATGATAAAGAGACAGTGATATTGG

FIGURE 34

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39518
<subunit 1 of 1, 440 aa, 1 stop
<MW: 48240, pI: 4.93, NX(S/T): 7
MASVVLPSPGSQCAAAAAAAPPGLRLLLLLFSAAALIPTGDGQNLFTKDVTVIEGEVATISCQ
VNKSDDSVIQLLNPNRQTIYFRDFRPLKDSRFQLLNFSSELKVSLTNVSISDEGRYFCQLYT
DPPQESYTTITVLVPPRNLMDIQKDTAVEGEEIEVNCTAMASKPATTIRWFKGNTELKGKSE
VEEWSDMYTVTSQMLMLKVHKEDDGVFVICQVEHPAVTGNLQTQRYLEVQYKPQVHIQMTYPLQ
GLTREGDALELTCEAIGKPQPMVTWVRVDDEMPQHAVLSGPNLFINNLNKTDNGTYRCEASN
IVGKAHSDYMLYVYDPPTTIPPTTTTTTTTTTTTTTILTIITDSRAGEEGSIRAVDHAVIGGV
VAVVVFAMLCLLIILGRYFARHKGTYFTHEAKGADDAADADTAIINAEGGQNNSEEKKEYFI

Signal sequence.

amino acids 1-36

Transmembrane domain.

amino acids 372-393

N-glycosylation sites.amino acids 65-69, 99-103, 111-115, 163-167, 302-306, 306-310,
430-434**Tyrosine kinase phosphorylation sites.**

amino acids 233-240, 319-328

N-myristoylation sites.amino acids 9-15, 227-233, 307-313, 365-371, 376-382, 402-408,
411-417, 427-433, 428-432

FIGURE 35

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCCTTGTGAGGAGGAGACAGCCTCCCGGCCCGGGGAG
 GACAAGTCGCTGCCACCTTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCTGCGGTGAGCTGCCGGCCG
 AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGCTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG
 AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
 TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGGAACACCTTTTGCTGTG
 ATGGCCCGCTCATGATGGCCCCGGCAAAAGGGCATTTTCTACCTGACCCTTTTCTCATCTCTGGGGACATGTACAC
 TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTGAGCTGTCTCCTGCCATCCCTGTATTGCTGCCATGCTCT
 TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCTCGGGCGCTACCAGATG
 AAGCAGCTTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
 TCAAGAATTTCCAGATAAACAACCAGATTGTGAACTGAAATACTGTTACACATGCAAGATCTTCGGCCCTCCCC
 GGGCCTCCCATTCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTG
 TTGGAAGAGGAACACCGCTACTTCTACCTCTCATCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCT
 TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAATTTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAAGTG
 TTCTAGAAGTCTCATTTGCTTCTTTACACTCTGGTCCGTCTGAGGACTGACTGGATTTTCACTTTTCTCGTGG
 CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
 ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA
 TTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
 CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
 AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTGTTGTGTT
 TAATTAGGGCTATGAGAGATTTGAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
 GTTTTTCTTTGGTCTTTAGTCACCCAGTTGACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT
 CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGGCCCTGGCACTGGTTCT
 CCATGGCCTCAGCCACAGGGTCCCCCTTGGACCCCCCTCTCTCCCTCCAGATCCAGCCCTCCTGCTTGGGGTCAC
 TGGTCTCATCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
 GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
 TGGGGTCAGAAGATTCTCTGGCCACCAAGTGCCAGCATTGCCACAAAATCCTTTTAGGAATGGGACAGGTACCT
 TCCACTTGTGTTGTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTGACTCCTGCTCCCATTAGGAG
 CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCTCAGAGGAAGCCCCGAGTGCTCACTTAAAC
 ACTATCCCCCTCAGACTCCCTGTGTGAGGCCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
 GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAAGCCT
 CGGCTGAGTGAGGGAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCCCACCTCTGGCT
 ATGGTAACCACACTGGGGGCTTCTTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCAGAGCCACTT
 CACCCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTTCAAGGAAGAAGATTTATGT
 ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
 GCGGTGGGGGAGTGTAACCGGAACTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

36/615

FIGURE 36

MSVMVVRKKVTRKWEKLPGRNTFCDDGRVMMARQKGIFYLTLFLIILGTCTLFFAFECRYLAVQ
LSPAIPVF AAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPPRI
KNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVVGKRNRYFYLFIL
SLSLLTIYVFAFNIVYVALKSLKIGFLETCLKETPGTVLEVLI CFFTLWSVVGLTGFHTFLVAL
NQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSRPPSTQ
ETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPEPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

1005336.01509

FIGURE 37

GGCGGAGCAGCCCTAGCCGCCACCGTCTGCTCTCGCAGCTCTCGTCCGCACTGCCACCGCCGCCGCCGCTCACTGCCG
TCCTGGCTCCGGCTCCCGCCCCCTCCCGGCCGGCCATGAGCCCCCGCCGCCAGGCGCCCGGTGCGCAGCTGC
TGCCCGCGCTGGCCCTGCTGCTGCTGCTGCTCGGAGCGGGGCCCCGAGGCAGCTCCCTGGCCAAACCCGGTGCCCG
CCGCGCCCTTGTCTGCGCCCCGGGCGGTGCGCCGCGCAGCCCTGCCGGAATGGGGGTGTGTGCACCTCGCGCCCTG
AGCCGGACCCGCAGCACCCGGCCCCCGCCGGCGAGCCTGGCTACAGCTGCACCTGCCCCGCCGGGATCTCCGGCG
CCAAGTCCAGCTTGTTCAGATCCTTGTGCCAGCAACCCCTTGTACCATGGCAACTGCAGCAGCAGCAGCAGCA
GCAGCAGCGATGGCTACCTCTGCATTTGCAATGAAGGCTATGAAGGTCCCAACTGTGAACAGGCACCTCCAGTC
TCCCAGCCACTGGCTGGACCGAATCCATGGCACCCTGACAGCTTCAGCCTGTTCTGCTACTCAGGAGCCTGACA
AAATCCTGCCTCGCTCTCAGGCAACGGTGACACTGCCTACCTGGCAGCCGAAAAACAGGGCAGAAAGTTGTAGAAA
TGAAATGGGATCAAGTGGAGGTGATCCAGATATTGCCTGTGGGAATGCCAGTTCTAACAGCTCTGCGGGTGGCC
GCCTGGTATCCTTTGAAGTGCCACAGAACACCTCAGTCAAGATTCCGGCAAGATGCCACTGCCTCACTGATTTTGC
TCTGGAAGGTACGGCCACAGGATTCCAACAGTGCCTCCCTCATAGATGGACGAAGTGTGACCCCCCTTCAGGCTT
CAGGGGGACTGGTCTCTCGGAGGAGATGCTCGCCTTGGGGAAATACTACTTTATTGGTTTGTGAATGATTTCTG
TGACTAAGTCTATTGTGGCTTTGCGCTTAACCTGGTGGTGAAGGTGAGCAGCTGTGTGCCGGGGGAGAGTCACG
CAAATGACTTGGAGTGTTTCAAGAAAAGGAAAATGCACCACGAAGCCGTGAGAGGCAACTTTTTCTGTACCTGTG
AGGAGCAGTACGTGGGTACTTTCTGTGAAGAATACGATGCTTGCCAGAGGAAACCTTGCCAAAACAACGCGAGCT
GTATTGATGCAAATGAAAAGCAAGATGGGAGCAATTTACCTGTGTTTGCCTTCTGGTTATACTGGAGAGCTTT
GCCAGTCCAAGATTGATTACTGCATCCTAGACCCATGCAGAAATGGAGCAACATGCATTTCCAGTCTCAGTGGAT
TCACCTGCCAGTGTCCAGAAGGATACTTCGATCTGCTTGTGAAGAAAAAGTGGACCCCTGCGCCTCGTCTCCGT
GCCAGAACAAACGGCACCTGCTATGTGGACGGGTACACTTTACCTGCAACTGCAGCCCGGGCTTCACAGGGCCGA
CCTGTGCCCCAGCTTATTGACTTCTGTGCCCTCAGCCCTGTGCTCATGGCACGTGCCGAGCGTGGGCACCACT
ACAAATGCCTCTGTGATCCAGGTTACCATGGCCTCTACTGTGAGGAGGAATATAATGAGTGCCTCTCCGCTCCAT
GCCTGAATGCAGCCACCTGCAGGGACCTCGTTAATGGCTATGAGTGTGTGTGCCCTGGCAGAATACAAAGGAACAC
ACTGTGAATTGTACAAGGATCCCTGCGCTAACGTGAGCTGTCTGAACGGAGCCACCTGTGACAGCGACGGCTGA
ATGGCACGTGCATCTGTGCACCCGGGTTTACAGGTGAAGAGTGGCAGATTGACATAAATGAATGTGACAGTAACC
CCTGCCACCATGGTGGGAGCTGCCTGGACCAGCCCAATGGTTATAACTGCCACTGCCCGCATGGTTGGGTGGGAG
CAAACGTGAGATCCACCTCCAATGGAAGTCCGGGCACATGGCGGAGAGCCTCACCAACATGCCACGGCACTCCC
TCTACATCATCATTTGGAGCCCTCTGCGTGGCCTTCATCCTTATGCTGATCATCCTGATCGTGGGGATTTGCCGCA
TCAGCCGCATTGAATACCAGGGTTCTTCCAGGCCAGCCTATGAGGAGTTCTACAAC TGCCGAGCATCGACAGCG
AGTTCAGCAATGCCATTGCATCCATCCGGCATGCCAGGTTTGGAAAGAAATCCCGGCTGCAATGTATGATGTGA
GCCCCATCGCCTATGAAGATTACAGTCTGATGACAAACCCCTTGGTCACACTGATTAAAACTAAAGATTTGTAAAT
CTTTTTTTGGATTATTTTTCAAAAAGATGAGATACTACACTCATTATAAATATTTTAAAGAAAATAAAAGCTTAA
GAAATTTAAATGCTAGCTGCTCAAGAGTTTTTCAGTAGAATATTTAAGAACATAATTTTCTGCAGCTTTTAGTTTG
GAAAAAATATTTTAAAAACAAAATTTGTGAAACCTATAGACGATGTTTTAATGTACCTTCAGCTCTCTAACTGT
GTGCTTCTACTAGTGTGTGCTCTTTTCACTGTAGACACTATCACGAGACCCAGATTAATTTCTGTGGTTGTTACA
GAATAAGTCTAATCAAGGAGAAGTTTCTGTTTTGACGTTTGAGTGCCGGCTTCTGAGTAGAGTTAGGAAAACAC
GTAACGTAGCATATGATGTATAATAGAGTATACCCGTTACTTAAAAAGAAGTCTGAAATGTTCTGTTTTGTGAAAA
AGAACTAGTTAAATTTACTATTTCCTAACCCGAATGAAATTAGCCTTTGCCTTATTCTGTGCATGGGTAAAGTAAC
TTATTTCTGCACTGTTTTGTGTAACCTTTGTGGAACATTCTTTTCGAGTTTGTTTTTGTGATTTTTCGTAAACAGTCG
TCGAACTAGGCCCTCAAAAACATACGTAACGAAAAGGCCTAGCGAGGCAAAATCTGATTGATTTGAATCTATATTT
TTCTTTAAAAAGTCAAGGGTTCTATATTGTGAGTAAATTAATTTACATTTGAGTTGTTTGTGCTAAGAGGTAG
TAAATGTAAGAGAGTACTGGTTCTTCAGTAGTGAGTATTTCTCATAGTGACGCTTTATTTATCTCCAGGATGTT
TTTGTGGCTGTATTTGATTGATATGTGCTTCTTCTGATTCTTGCTAATTTCCAACCATATTGAATAAATGTGATC
AAGTCA

FIGURE 38

><subunit 1 of 1, 737 aa, 1 stop
 ><MW: 78475, pI: 5.09, NX(S/T): 11
 MQPRRAQAPGAQLLPALALLLLLLLGAGPRGSSSLANPVPAAPLSAPGPCAAQPCRNGGVCTSRP
 EPDPQHPAPAGEPGYSCTCPAGISGANQQLVADPCASNPCHHGNCSSSSSSSSSDGYLCICNEG
 YEGPNCEQALPSLPATGWTESMAPRQLQPVATQEPDKILPRSQATVTLPTWQPKTGQKVDEM
 KWDQVEVIPDIACGNASSNSSAGGRLVSFEVPQNTSVKIRQDATASLILLWKVTATGFFQQCSL
 IDGRSVTPLQASGGLVLLLEEMLALGNHFIGFVNDVSVTKSIVALRLTLVVKVSTCVPGESHAN
 DLECSGKGKCTTKPSEATFSCCTCEEQYVGTFCEEYDACQRKPCQNNASCIDANEKQDGSNFTC
 VCLPGYTGELCQSKIDYCILDPNRGATCISSLSGFTCCQCEGYFGSACEEKVDPCASSPCQN
 NGTCYVDGVHFTCNCSPGFTGPTCAQLIDFCALSPCAHGTCRSVGTSYKCLCDPGYHGLYCEE
 EYNECLSAPCLNAAATCRDLVNGYECVCLAEYKGTCELYKDPCANVSCNLGATCDSDGLNGTC
 ICAPGFTGEECDIDINECDSNPCHHGGSCLDQPNGYNCHCPHGWVGANCEIHLQWKSGHMAES
 LTNMPRHSLYIIIGALCVAFILMLIILIVGICRISRIEYQGSSRPAYEEFYNCRSIDSEFSNA
 IASIRHARFGKKSRRPAMYDVSPYAYEDYSPDDKPLVTLIKTKDL

Signal sequence.

amino acids 1-28

Transmembrane domain.

amino acids 641-660

N-glycosylation sites.

amino acids 107-111, 204-208, 208-212, 223-227, 286-290, 361-365,
 375-379, 442-446, 549-553, 564-568

Glycosaminoglycan attachment site.

amino acids 320-324

Tyrosine kinase phosphorylation sites.

amino acids 490-498, 674-682

N-myristoylation sites.

amino acids 30-36, 56-62, 57-63, 85-91, 106-112, 203-209,
 373-379, 449-455, 480-486, 562-568, 565-571

Amidation site.

amino acids 702-706

Aspartic acid and asparagine hydroxylation site.

amino acids 520-532, 596-608

EGF-like domain cysteine pattern signatures.

amino acids 80-92, 121-133, 336-348, 378-390, 416-428, 454-466,
 491-503, 529-541, 567-579, 605-617

FIGURE 39

GAGCCGCCGCCGCGCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCGGCCCGCCCCACCCACGTCTG
 CGTTGCTGCCCCGCCTGGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACTCC
 GCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGGTTTGCAGGGGGCT
 TCGCGTGTTCGAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTGCTGC
 TAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGCTCGGCG
 TGGTCATTGCAGTGGGCATCTTCTTGTTTCTGATTGCTTTAGTGGGTCTGATTGGAGCTGTAA
 AACATCATCAGGTGTTGCTATTTTTTTATATGATTATTCTGTTACTTGTATTTATTGTTTCAGT
 TTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCTGGAGGTTG
 GTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCTGTGGGTTCC
 GAAGTGTTAACCCTAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCACTCGTGCTCGC
 CATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGGTGGCATTTGGCC
 TGTTCCTTCAGTTTTTACAGAGATCCTGGGTGTTTTGGCTGACCTACAGATACAGGAACCAGAAAG
 ACCCCCGCGCGAATCCTAGTGCAATTCCTTTTGATGAGAAAACAAGGAAGATTTCTTTTCGTATT
 ATGATCTTGTTCACTTTCTGTAATTTCTGTTAAGCTCCATTTGCCAGTTTAAGGAAGGAAAC
 ACTATCTGGAAAAGTACCTTATTGATAGTGGAATTATATATTTTTACTCTATGTTTCTCTACA
 TGTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTCTGAAGCTCGGTGGCAC
 CTGGAATTTACTGTATTCATTGTGCGGCACTGTCCACTGTGGCCTTTCTTAGCATTTTTACCT
 GCAGAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAATCTGAACGTACATCTCAC
 TGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCCTACTGGAAAAAGAGTGGAATT
 TATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAATCCAAATTCCTAATTTTTTT
 TGGTCTTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTATAAAAAATGATAATTTACTTGT
 AGTCTTTTTATGATTACACCAATGTATTCTAGAAATAGTTATGTCTTAGGAAATGTGGTTTTAA
 TTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGTTTTCATGAAATGTTCTAATGTATAAT
 AACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTTTGAGTAATCAGGAAGTATATCTAT
 ATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAAAAGACTGCATTTTTTAAACAAGTTAG
 TATTAATGCGTTGGCCACGTAGCAAAAAGATATTTGATTATCTTAAAAATTGTTAAATACCG
 TTTTCATGAAATTTCTCAGTATTGTAACAGCAACTTGTCAAACCTAAGCATATTTGAATATGA
 TCTCCCATAATTTGAAATTGAAATCGTATTGTGTGGCTCTGTATATTCTGTTAAAAAATTAA
 GGACAGAAACCTTTCTTTGTGTATGCATGTTTGAATTAAAAGAAAGTAATGGAAG

40/615

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979
><subunit 1 of 1, 204 aa, 1 stop
><MW: 22147, pI: 8.37, NX(S/T): 3
MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALVG
LIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQRNL
NCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFRVGGIGLFFSFTEILGVWLTYR
YRNQKDPRANPSAFL
```

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

40/615

FIGURE 41

CAGTCACCA**ATGA**AGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTGC
TCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGTCT
GCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCTTCC
AGGTCAAGGCCTACACTTTTCAGTGAACCTTCCACCTGATTGTGTCTATGACTGGCTGATCC
TCCAAGGTCCAGCCAAGCCAGTTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGCCTGGC
AAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCCCGGGC
CTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCACTGCAGTG
GCATCTTCCAGAGCCCTGGTCCTGGGATCCCAGAAACAGCATCTGTTGTGGCTATCACAGTCC
AAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAGCAGGAAGCC
CCATGACCCTGAGTTGTAGACAAAAGTTGCCCTGCAGAGGTCAGCTGCCCCGCTCCTCTTCT
CCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATTCAGATCCCCA
CAGCTTCAGAAGATCACTCCGGGTCACTACTGGTGTGAGGCAGCCACTGAGGACAACCAAGTTT
GGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCTGCTGCACCTCCCA
CATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGAGGAGGCCCTGGGC
CTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTTCTCCTCTGGGGATGC
CAGATCCTCATCTGTATCACCGATGGGCCTTCTTCTCAAACACATGCAGGATGTGAGAGTCC
TCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCAGAAGCCTGGGACCACAA
AGGCTACTGCTGAAT**AGAA**AGTAAACAGTTTCATCCATGATCTCACTTAACCAACCCCAATAAATC
TGATTCTTTATTTTCTCTTCCTGTCTGCACATATGCATAAGTACTTTTACAAGTTGTCCCAG
TGTTTTGTTAGATAATGTAGTTAGGTGAGTGTAATAAATTTATATAAAGTGAGAATTAGAG
TTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTCTGCTGTCTAGATCAGGAATTT
CTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAACTAATGGAAGTGGATTGAATAC
AGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACATTGGGCAATGTTTGGAGACATTT
TGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGTGTGTCTACTGGCATCCAGTAAATA
GAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAGGGCAGTACCCCAACGAAAAATAA
TCTGGCCCCAAAATGTCAGTTGTACTGAGTTTGAGAAACCCAGCCTAATGAAACCTAGGTGT
TGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTATTATCTCTTTCCAGCCTCATTGAGCTAT
TCTTACTGACATAACAGTCTTTAGCTGGTGCTATGGTCTGTTCTTTAGTTCTAGTTTGTATCC
CCTCAAAAGCCATTATGTTGAAATCCTAATCCCCAAGGTGATGGCATTAGAAGTGGGCCTTT
GGGAAGTGATTAGATCAGGAGTGCAGAGCCCTCATGATTAGGATTAGTGCCCTTATTTAAAAA
GGCCCCAGAGAGCTAACTCACCTTCCACCATATGAGGACGTGGCAAGAAGATGACATGTATG
AGAACCACAAAAACAGCTGTCGCCAAACACCGACTCTGTCGTTGCCTTGATCTTGAACCTCCAG
CCTCCAGAACTATGAGAAATAAAATTCTGGTTGTTTGTAGCCTAA

42/615

FIGURE 42

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLGCVLMAWALYLSLGVLWVAQMMLLAASFETLQCEGPVCTEESCHTEDDLTDAREAGFQVK
AYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQAWQDWPLTQVTFYRDGSALGPPGPNR
EFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGSPMT
LSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQVWKQ
SPQLEIRVQGASSAAPTLPNPAPQKSAAPGTAPEEAPGPLPPPPTPSSDPGFSSPLGMPDP
HLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 43

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTTC
TGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGGAA
GGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAACTCTGGGTGTCTGCATT
GCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTGGGCA
CATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATCCTTGT
GGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGACTAGCAA
GTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGGTGAAGCG
GCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTCTGCAGGACTCAGCTGCAGGTTTTAT
CGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGGAGCTGCCAA
AGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCAGAGGGGAACT
TCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATGGGCCGCTCGGC
CTACAATGAAGGGGACTATTATCATACGGTGTTGTGGATGGAGCAGGTGCTAAAGCAGCTTGA
TGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCAGCTATGCTGTCTT
CCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCTGCTCTCCCTTGACCCAAG
CCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGAGGAAGAGAGAGAAAA
AACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAGGCATCTATGAGAGGCCTGT
GGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGTTGGGGAGGGTGTCAAACCTGAC
ACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAACAGGGCCCCACAGCTGCT
CATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACATCGTCAGGTACTACGATGT
CATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAACCTAAACTTGCACGAGCCAC
CGTTCGTGATCCCAAGACAGGAGTCCCTCACTGTGCCAGCTACCGGGTTTTCCAAAAGCTCCTG
GCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATCGTCGGATGCAGCATATCACAGG
GTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAATTATGGAGTGGGAGGACAGTATGA
ACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGGCCTCAAAACAGAGGGGAATAGGTT
AGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTGGTGGTGCCACCGTCTTCCCTGATCT
GGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTGTTCTGGTACAACCTCTTGCGGAGCGG
GGAAGGTGACTACCGAACAAGACATGCTGCCTGCCCTGTGCTTGTGGGCTGCAAGTGGGTCTC
CAATAAGTGGTTCCATGAACGAGGACAGGAGTTCTTGAGACCTTGTGGATCAACAGAAGTTGA
CTGACATCCTTTTTCTGTCTTCCCCCTTCCCTGGTCTTTCAGCCCATGTCAACGTGACAGACACC
TTTGTATGTTCTTTGTATGTTCCCTATCAGGCTGATTTTTGGAGAAATGAATGTTTGTCTGGA
GCAGAGGGAGACCATACTAGGGCGACTCCTGTGTGACTGAAGTCCCAGCCCTTCCATTACGCC
TGTGCCATCCCTGGCCCCAAGGCTAGGATCAAAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGC
CTAGCAAGGTGCCTTTGTACCTCAGGTGTTTTAGGTGTGAGATGTTTCAGTGAACCAAAGTTC
TGATACCTTGTTTACATGTTTGTGTTTTATGGCATTTCTATCTATTGTGGCTTACCAAAAAAT
AAAAATGTCCCTACCAGAAAAA

FIGURE 44

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWAN
KMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFPTD
EDEIGAALKMRLQDTYRLDPGTISRGEPLPGTKYQAMLSVDDCFGMGRSAYNEGDDYYHTVLWM
EQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLRYFEQ
LLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEVGLTPRRQKRLFCRYHH
GNRAPQLLIAPFKEEDEWDSPHIVRYDVMSEDEIERIKEIAKPKLARATVRDPKTGVLTVAS
YRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFSRRPFDG
LKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYRTRHAACP
LVGCKWVSNKWFHERGQEFRLPCGSTVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534**Tyrosine kinase phosphorylation site.**

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 45

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGGG
 CGGCGGGCGCGGGTGCAGGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCTCGCTCCCAGCC
 TGTCTGTCTGTCGTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTCACACCCGATCCTGGGC
 TTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAGCGGG
 TCGTCTGTGTCTCTCTCTCTCTGCGCCGCGCCGGGGATCCGAAGGGTGCAGGGGCTCTGAGGA
 GGTGACGCGCGGGGCCTCCCGCACCCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAGGTGTGA
 GCAGCCTATCAGTCACCA**ATGT**CCGCGAGCCTGGATCCCGGCTCTCGGCCTCGGTGTGTGTCTGC
 TGCTGCTGCCGGGGCCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATGTTTTACCA
 GAGGCTTGACATCAGGAAAGAGAAAGCAGATGTCTCTGCCCAGGGGGCTGCCCTCTTGAGG
 AATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGGGCTGCTGTCC
 ACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACCTGGTTCGAGAAA
 ACTATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGATGGTCTGCTTCTT
 TCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCAGTGTCCACAGCAC
 ATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGGCAATAAAGATTGTA
 AAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGCGCCGATTTAATTTAC
 AGAAGAATTTTGTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACAGAAGGACCACATGTGG
 GCCTTGTTCAAGCCAGTGAACATCCCAAATAGAAATTTTACTTGAAAACTTTACATCAGCCA
 AAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTTAGAGGGGGTAATTCCAATACAGGAAAAG
 CTTTGAAGCATACTGCTCAGAAATTCCTCACGGTAGATGCTGGAGTAAGAAAAGGGATCCCCA
 AAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACATCGAGGAAGCAGGCATTGTGG
 CCAGAGAGTTTGGTGTCAATGTATTTATAGTTTTCTGTGGCCAAGCCTATCCCTGAAGAACTGG
 GGATGGTTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGTGCGGAATAATGGCTTCTTCTCTT
 ACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAAGCCTCTGGTACAGAAGCTGTGCA
 CTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACTCAGTGAACATTGCCTTTCTAATTG
 ATGGCTCCAGCAGTGTGGAGATAGCAATTTCCGCCTCATGCTTGAATTTGTTTCCAACATAG
 CCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGATAGCTGCTGTACAGTTTACTTATGATC
 AGCGCACGGAGTTCAGTTTCACTGACTATAGCACCAAAGAGAATGTCTTAGCTGTATCAGAA
 ACATCCGCTATATGAGTGGTGGAACAGCTACTGGTGATGCCATTTCTTCACTGTTAGAAATG
 TGTTTGGCCCTATAAGGGAGAGCCCCAACAAAGAACTTCCTAGTAATTGTACAGATGGGCAGT
 CCTATGATGATGTCCAAGGCCCTGCAGCTGCTGCACATGATGCAGGAATCACTATCTTCTCTG
 TTGGTGTGGCTTGGGCACCTCTGGATGACCTGAAAGATATGGCTTCTAAACCGAAGGAGTCTC
 ACGCTTTCTTCACAAGAGAGTTCACAGGATTAGAACCAATTGTTTCTGATGTATCAGAGGCA
 TTTGTAGAGATTTCTTAGAATCCCAGCAAT**TAA**TGGTAACATTTTGACAACTGAAAGAAAAAGT
 ACAAGGGGATCCAGTGTGTAAATTGTATTCTCATAATACTGAAATGCTTTAGCATACTAGAAT
 CAGATACAAAATATTAAGTATGTCAACAGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCG
 CTGCCTTCTGGTTACAATTTACAGTGTACTTTGTTAAAAACACTGCTGAGGCTTCATAATCAT
 GGCTCTTAGAAACTCAGGAAAAGAGGAGATAATGTGGATTAAACCTTAAGAGTTCTAACCATG
 CCTACTAAATGTACAGATATGCAAATTCATAGCTCAATAAAAGAATCTGATACTTAGACCAA
 AAAAAAAA

46/615

FIGURE 46

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEF SVYG
NIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVTKG
KSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKNFVG
KVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAlKEVGFRGGNSNTGKALKHTA
QKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELGMVQDV
TFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNI AFLIDGSSSV
GDSNFRMLLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVIRNIRYMS
GGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITIFSVGVAVA
PLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

47/615

FIGURE 47

CCCCGCGCCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAG
CCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCT
GCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGG
GGGTGCTGGTGTCCATCATTATGCTGAGCCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCT
GGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCC
TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTT
TCACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGT
TTTGTTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACG
GCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCATCCTCATCC
AGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGG
AGTGCGATTCCCGTGCCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGT
CGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCCAGCGGCTGCCACGAGGGCA
AGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTTGCCCA
AGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCA
TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAA
CCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATG
CCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTTCATCAGTCTGCGCTCCT
CAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCA
CACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACG
GCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCTGGCCTCACTGCACGTCATGA
TGACGCTCACCAACTGGTACAAGCCCCGGTGAGACCCGGAAGATGATCAGCACGTGGACCGCCG
TGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCTACCTGTGGACCCTGGTAGCCC
CACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTCACAGCCTGCCATCTGGTGCCTC
CTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTCCCCACACCAATCAGCC
AGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCCGGGCCTTCTAGTCGT
AGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCACACCCAC
ACGGTGGAGCTGCCTCTTCCTTCCCCCTCCTCCCTGTTGCCCATACTCAGCATCTCGGATGAAA
GGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACTCCCACC
ACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGACCCTGCC
CCCTTCCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAACAAGCCAGTGCG
TGTAACAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 48

MGACLGACSLLSASCCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLSIIMLSPGVES
QLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPR
AAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQR
WLGKAEEDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVS
IAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGY
ETQWWDAPSIIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRA
FDNEQDGVITYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLY
LWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

[illegible]

50/615

FIGURE 50

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVSS
SCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKSRR
KRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKGTQK
LRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPGWIKGNANDIGMDYDYALLEL
KKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDILLYQQCDAQPGA
SGSGVYVRMWKRQQQKWERKIIGIFSGHQWDMNGSPQDFNVAVRITPLKYAQICYWIKGNYL
DCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 51

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG
 CAGTAGCCCGGGCGGCGAGGGCTGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCACCACCC
 CAACCTGTTTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAAC**ATGG**ATTTTCTCCTGGCGCTGGT
 GCTGGTATCCTCGCTCTACCTGCAGGCGGCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGAT
 TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC
 TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG
 AAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA
 CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG
 CTCCATGGCAAACCTGTGAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCCTGGCCT
 GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT
 TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTGATAAAGGCTTCGATCTCATGTATATTGGAGG
 CAAATATCAATGTATGACATAGACGAATGCTCATTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
 CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA
 AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA
 TAATAATTGGATTCTCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCTCCTATCATTACCAA
 CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCTACTCCACCACCACCACCACC
 CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAAGGCAACACCCGGACTGACAACCTATAGC
 ACCAGCTGCCAGTACACCTCCAGGAGGATTACAGTTGACAAACAGGGTACAGACAGACCCCTCAGAAACCCAGAGG
 AGATGTGTTTCACTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA
 TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTGCGCAGCCAAAGCCCCAGG
 GGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCTCATGCAATTCAGGGGACCTGTGCCCTGTCATTTCAGGCA
 CAAGGTGACGGGGCTGCACCTCTGGCACACTCCAGGTGTTTGTGAGAAAAACAGGTGCCACGGAGCAGCCCTGTG
 GGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA
 AAGAT**GA**TTAAAGGGTTGGAAAAAAGATCTATGATGGAAAAATTAAGGAACCTGGGATTATTGAGCCTGGAGAAG
 AGAAGACTGAGGGGCAAACATTGATGGTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG
 TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG
 GGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTCTTTCTAAAAAATTAGA
 TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAAGTAACAAATTTATAGAATTTCCCAAA
 AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG
 CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCAACAAATGCTGAGCTCACTGAAATA
 TCTCTCCCTTATGGCAATCCTAGCAGTATTAAGAAAAAAGGAACTATTTATTTCCAAATGAGAGTATGATGGAC
 AGATATTTTAGTATCTCAGTAATGTCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC
 TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTTTAAAGAGATCCTTCAAGGAACACAGTTCAGAGAG
 ATTTTCATCGGGTGCAATCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAAGAGTGCCCTGCCCC
 ACACCGGCAGACCTTTCTTTCACCTCATCAGTATGATTCACTTCTTATCAATTGGACTCTCCAGGTTCCAC
 AGAACAGTAATATTTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTTAACCTGGTAAAGGCAGGGCTGG
 AGGGGGAAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTAAATGGTTCATT
 TCCTTTATGGTCAATATACTGCACAGCTGAAGATGAAAGGGGAAAAATAAATGAAAATTTTACTTTTCGATGCCAA
 TGATACATTGCACTAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTATTATTATTAATGTTTT
 CTAATAAAAAATGTTAGTGGTTTTCCAAATGGCCTAATAAAAAACAATTATTTGTAAATAAAAAACACTGTTAGTAAT

FIGURE 52

MDFLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQPR
CKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGS
CSSALTCSMANCQYGCDVVKGQIRCQCPSPLHLAPDGRTCVDVDECATGRASCPRFRQCVNT
FGSYICKCHKGFDLMIYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLT
CVYIPKVMIEPSGPIHVPKGNLTILKQDGTGNNNWIPDVGSTWWPPKTPYIIPPIITNRPTSKPT
TRPTPKPTPIPTPPPPPLPTELRTPLPPTTTPERTTGLTTIAPAASTPPGGITVDNRVQTD
QKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVL
PLGRLMHSGDLCLSFHVKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIK
SESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482**Aspartic acid and asparagine hydroxylation site.**

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 53

CGGGCCGCCCCCGGCCCCCATTCGGGCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGGG
 CCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAGCC
 CAGCTGTGAGCCGCCTCACAGGAAG**ATG**CTGCGTCGGCGGGGCAGCCCTGGCATGGGTGTGCA
 TGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGGTCCC
 TGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCCCCTGA
 GCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATAACCAAACAGCTGGT
 GCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCCTCTTCCC
 GGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCGGACGAGGG
 CAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCCGGCAGCGCTGCCGTGAGCCTGCAGGTGGC
 CGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGCCAGGGGACAC
 GGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTCTGGCAGGATGG
 GCAGGGTGTGCCCCTGA CTGGCAACGTGACCACGTGCGAGATGGCCAACGAGCAGGGCTTGTT
 TGATGTGCACAGCGTCCTGCGGGTGGTGTGGGTGCGAATGGCACCTACAGCTGCCTGGTGCG
 CAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTACCATCACAGGGCAGCCTATGACATT
 CCCCCAGAGGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCATTGCACTGCTGGTGGC
 CCTGGCTTTCTGTGTGCTGGAGAAAGATCAAAACAGAGCTGTGAGGAGGAGAATGCAGGAGCTGA
 GGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCTCTGAAACACTCTGACAG
 CAAAGAAGATGATGGACAAGAAATAGCC**TGA**CCATGAGGACCAGGGAGCTGCTACCCCTCCCT
 ACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCCTGCCCCCAACAGATGCATC
 CTGCTCTGACAGGTGGGCTCCTTCTCCAAGGATGCGATACACAGACCACTGTGCAGCCTTAT
 TTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTTTTTCTTATAGACACAATGAACA
 GACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCTGCCTTATTTTACAGTACATACA
 TTTCTTAGGGACACAGTACACTGACCACATCACCACCCTCTTCTTCCAGTGCTGCGTGGACCA
 TCTGGCTGCCTTTTTTTCTCCAAAAGATGCAATATTCAGACTGACTGACCCCTGCCTTATTTT
 ACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTCTCCAATGGCCGTGATACACTAGTGA
 TCATGTTTCAGCCCTGCTTCCACCTGCATAGAATCTTTTCTTCTCAGACAGGGACAGTGCGGCC
 TCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTTCCCCTCCTTCTCCCTGCCCCAAGTGAA
 GACAGGGCAGGGCCAGGAATGCTTTGGGGACACCGAGGGGACTGCCCCCACCCCCACCATGG
 TGCTATTCTGGGGCTGGGGCAGTCTTTTCTGGCTTGCTCTGGCCAGCTCCTGGCCTCTGGT
 AGAGTGAGACTTCAGACGTTCTGATGCCTTCCGGATGTCATCTCTCCTGCCCCAGGAATGGA
 AGATGTGAGGACTTCTAATTTAAATGTGGGACTCGGAGGGATTTTGTAACCTGGGGGTATATT
 TTGGGGAAAATAAATGTCTTTGTAAAAAAAAAAAAAAAAAAAAA

54/615

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQL
NLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVADEGSFTCFVSI
RDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDQGQGVPLTGN
VTTSQMANEQGLFDVHSLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEALWVT
VGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEKSKTALQPLKHSDSKEDDGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

20250610 10:52:00

FIGURE 55

GAGTCTTGACCGCCGCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCGCTG
GCT**TATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTT
CTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAG
TGTGACCACGTGCAATATACGCTGGTTCAGTTTCTGGGTGGCAAGAACTTGAAACTGCATTT
CTTGAGCATAAAGAACAGTTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTA
TTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGGCCAGTC
AATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGATGACCTT
GAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATTTCAGGAAAT
GACAGTGATGGGTTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATAGTGGAGCAA
ACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCTCTTTGACTAC
GAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGGCTTGGATGCTG
TCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGACCAGTGGGTGCAA
GACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTTGCAGCGCCACGTTTCCCGC
CACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCACACGGATCTCCTTT
GAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGACAGCCTGTGCAACACC
AGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAAGCGGCTCCAGGAGTTC
CTTGACAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCCAGGCCATGGACATCTCC
TTGAAGGAGAAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAATTTGGGATGAAGGACATG
CGCGTGCAGACTTTTCAGCATTCATTTTGGGTTCGAAGCACAAAGTTTCTGGCCAGCGACGTGGTC
TTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCTCAGGGACAGATCACTTCATC
CAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTGTACCATGGCCTGGAACCTCGCC
AAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGCCTTTGCACCAACCTCGTCATC
TCCCAGGGGGCCTTTCTGTACTGCTCTCTCATGGAGGGCACTCCAGATGTCATGCTGTTCTCT
AGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCAAGTCCTTTGTGTGTTTCGACAAAG
AACC GGCGCTGCAAACTGCTGCCCCCTGGTGATGGCTGCCCCCTGAGCATGGAGCATGGCACA
GTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTCGGACAGGAAGAACTTTTTTGGGAGG
GCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGATGCTGCACAACCATTTTGACCTCTCA
GTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTTCTGGACGCACTTATTTCCCTCCTGTCC
TAGGAATTTGATTCTTCCAGAATGACCTTCTTATTTATGTAACCTGGCTTTTCATTTAGATTGTA
AGTTATGGACATGATTTGAGATGTAGAAGCCATTTTTTATTAAATAAAATGCTTATTTTAGGAAA

56/615

FIGURE 56

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
EHKEQFHYFILINCGANVDLLDILQPDEDTIFVCDSHRPVNVVNVYNDTQIKLLIKQDDDDLE
VPAYEDIFRDEEEDEEHSGNDSGSEPSEKTRLEEEIVEQTMRRRQRREWEARRRDILFDYE
QYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLRHVSRH
NHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKRLQEFL
ADMGLPLKQVKQKFQAMDISLKENLREMIIEESANKFGMKDMRVQTF SIHFGFKHKFLASDVVF
ATMSLMESPEKDGS GTDHF IQALDSLRSNLDKLYHGLELAKKQLRATQQT IASCLCTNLVIS
QGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKS FVCSTKNRRCKLLPLVMAAPLSMEHGT
TVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFLDALISLLS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

57/615

FIGURE 57

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGGC
CCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTTCGCCTTGCTGACGGCGTCGAG
CCCTGGCCAGACATGTCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACCGTG
GCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAACCCT
TCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTGCTCCT
TCAAGTGGTTTTTGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTAGGAGGA
ACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAACCCTGCAA
GGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCTTCTCCAGA
CCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGAGCCCTGGAAAGGAATCAGA
GATGCTACCACCTACCCGCTGGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCA
CGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGCCTCAGCCTCTG
AGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCTCGATGTACGTACGC
ACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCTGAACGTGTACGCG
CCGGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCCTGGGAGGCGCCTTC
ATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCCGCGAGAAAGTGGTGCTG
GTGTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCTTGAGCACGGACGACAGCCACGCGCGC
GGAACTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGCAGGAGAACATCGCAGCC
TTCGGGGGAGACCCAGGAAATGTGACCCTGTTTCGGCCAGTCGGCGGGGGCCATGAGCATCTCA
GGACTGATGATGTCACCCCTAGCCTCGGGTCTCTTCATCGGGCCATTTCCAGAGTGGCACC
GCGTTATTACAGACTTTTCATCACTAGTAACCCACTGAAAGTGGCCAAGAAGGTTGCCACCTG
GCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGCCTGAGGGGCACTATCAGGGACC
AAGGTGATGCGTGTGTCCAACAAGATGAGATTCCTCCAACCTGAACCTCCAGAGAGACCCGGAA
GAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGGTGATCCAGATGACCCTTTGGTG
CTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTTCTAGGTGTCAACAACCTGGAATTC
AATTGGCTCTTGCCTTATAATATACCAAGGAGCAGGTACCACTTGTGGTGGAGGAGTACCTG
GACAATGTCAATGAGCATGACTGGAAGATGCTACGAAACCGTATGATGGACATAGTTCAAGAT
GCCACTTTCGTGTATGCCACACTGCAGACTGCTCACTACCACCGAGAAACCCCAATGATGGGA
ATCTGCCCTGCTGGCCACGCTACAACAAGGATGAAAAGTACCTGCAGCTGGATTTTACCACAA
GAGTGGGCA~~TGA~~AGCTCAAGGAGAAGAAGATGGCTTTTTTGGATGAGTCTGTACCAGTCTCAAA
GACCTGAGAAGCAGAGGCAATTCTAAGGTGGCTATGCAGGAAGGAGCCAAAGAGGGGTTTGC
CCCCACCATCCAGGCCCTGGGGAGACTAGCCATGGACATACCTGGGGACAAGAGTTCTACCCA
CCCCAGTTTGAAGTGCAGGAGCTCCCTGCTGCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGT
TGTGTGGGACCTGCACTGCCCTTTCCAGCCTGACATCCCATGATGCCCCTCTACTTCACTGTT
GACATCCAGTTAGGCCAGGCCCTGTCAACACCACACTGTGCTCAGCTCTCCAGCCTCAGGACA
ACCTCTTTTTTTCCCTTCTTCAAATCCTCCACCCCTTCAATGTCTCCTTGTGACTCCTTCTTA
TGGGAGGTGACCCAGACTGCCACTGCCCTGTCACTGCACCCAGCTTGGCATTTACCATCCA
TCCTGCTCAACCTTGTTCCCTGTCTGTTCAATTGGCCTGGAGGCCTAGGGCAGGTTGTGACAT
GGAGCAAACCTTTTGGTAGTTTGGGATCTTCTCTCCACCCACACTTATCTCCCCAGGGCCAC
TCCAAAGTCTATACACAGGGGTGGTCTCTTCAATAAAGAAGTGTGATTAGAAAAA

2025-10-23 10:44:00

58/615

FIGURE 58

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSAPSSGF
GTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQKQMHVGKTPIQVFLGVFPFSRPPLG
ILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLSVWG
YRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAFIVGA
ASSYEGSDLAAREKVVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIAAFGGD
PGNVTILFGQSAGAMSIISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVAHLACN
HNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQORDPEEIIWSMSPVVDGVVIPDDPLVLLTQ
GKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMDIVQDATFV
YATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

60/615

FIGURE 60

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44192
<subunit 1 of 1, 694 aa, 1 stop
<MW: 77400, pI: 9.54, NX(S/T): 6
MLLLLGLCLGLSLCVGSQEEAQSWGHSSEQDGLRVPRQVRLLQRLKTKPLMTEFSVKSTIISR
YAFTTVSCRMLNRASEDQDIEFQMQUIPAAAFITNFTMLIGDKVYQGEITEREKKSGDRVKEKR
NKTTEENGEKGTEIFRASAVIPSKDKAAFFLSYEELLQRRLLGKYEHSISVRPQQLSGRLSVDV
NILESAGIASLEVLPLHNSRQRGSGRGEDDSGPPPSTVINQNETFANIIFKPTVVQQARIAQN
GILGDFIIRYDVNREQSIGDIIQVLNGYFVHYFAPKDLPLPKNVVFLDSSASMVGTKLRQTK
DALFTILHDLRPQDRFSIIIGFSNRIKVWKDHLISVTPDSIRDGKVYIHHMSPTGGTDINGALQ
RAIRLLNKYVAHSGIGDRSVSLIVFLTDGKPTVGETHTLKIILNNTREAARGQVCIFTIGIGND
VDFRLLLEKLSLENCGLTRRVHEEEDAGSQLIGFYDEIRTPLLSDIRIDYPPSSVVQATKTLFP
NYFNGSEIIIIAGKLVDRLDHLHVEVTASNSKKFIIILKTDVPVRPQKAGKDVGTGSPRPGGDGE
GDTNHIERLWSYLTTKELLSSWLQSDDEPEKERLRQRAQALAVSYRFLTPFTSMKLRGPVPRM
DGLEEAHGMSAAMGPEPVVQSVRGAGTQPGPLLKKPNSVKKKQNKTKKRHGRDGVFPLHHLGIR

Signal sequence.

amino acids 1-14

N-glycosylation sites.

amino acids 97-101, 127-131, 231-235, 421-425, 508-512, 674-678

Glycosaminoglycan attachment sites.

amino acids 213-217, 391-395

N-myristoylation sites.

amino acids 6-12, 10-16, 212-218, 370-376, 632-638, 638-644

ACGGAACCCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
ACGAAGGGGACGATACCAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTGGC
TGACTTTTGGCTATAGAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAAGTC
TTCCCTTTTATCGAGTACAAGAAACCCCTTCTTGAGCTATTACAGCTTTTAAACATTGAGT
AAAGTACGCTCCGGTCAACC**ATGGT**GACAGCGCCCTGGGTCCCGTCTGGGCAGCGCTCCTGCT
CTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGCCAGCGG
CTGCCAACGGTGCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAGCCTCTTC
CTCCGGCCGCCCCACGCCCCTGCTGAGATCAGACCCCTACATTAATATACCATTCTCTGAAGGG
TGACAAAGGGGACCCAGGCCCAATGGGCTCTGCCAGGGTACATGGGCAGGGAGGGTCCCCAAGG
GGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCGGCGCCCCGTG
CCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCGCAAGACGGCCCTGCACAGCGGCGAGGACTT
CCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGACATGGCGACCGG
CCAGTTTGCTGCTCCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGCACAGCTTGAATTA
CAAGGAGACGTACGTGCACATTATGCATAACGAGAAGAGGCTGTATCCTGTACGCGCAGCC
CAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGCTTGGCTACGGGGACCGCT
CTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCAACGACTTCGACACCTA
CATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACT**TGAG**GGGCCTCTGGGCCACCCTCC
CGGCTGGAGAGCTCAGGTGCTGGTCCCGTCCCTTGCAAGGCTCAGTTTGACATCTGTGTAAGC
AGGAAGGCCAGGGAGGTCCCCGGGACCTGGCATTCTGGGGAGACCCCTGCTTCTATCTTGGCT
GCCATCATCCCTCCAGCCTATTTCTGCTCCTCTCTTCTCTCTTGGACCTATTTTAAAGAAGCT
TGCTAACCTAAATATTTCTAGAACTTTCCAGCCTCGTAGCCCAGCACTTCTCAAACCTTGGAAA
TGCATGCGAATCACCCGGGGTTTCGTGTTAAATGCAGATTCTGACTCAGCAGGTCTGAGTGGGT
CCAGGATTCTGTGTTTCTCATATGTTCTCTGGGTGATGCTGATGGGGTCACTATGAACCACA
CTGGAGCAACAGGTTCTAGAGATTTCTCAATATTTCTAGTACTTTCTGAACATTCTGGAATCC
TCCCCACATTTCTAGTAATTTCTCCCAACATTTTTTTTTTCTTGAGACAGAGTCTTGCTCTGTGCG
CAGGCTAGAGTGCAGTGGTGCAATCTCAGTTCACCTGCAACCTCTGCCTCCCGGGTTCAAGCGA
TTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTACAGGCGCCTGCTACCATGCCTGGCTAAT
TTTTGTATTTTATGATAGAGATGGGGTTTCAACCATATTGGCCAGGCTGGTCTTGAACCTCTGAC
TTCAGTGTACCCACCGCTCTGGCTCTCAAATGCTGGGATTACAGGTGTGAGCCACCGTGC
CTGGCCAATTTCAAACATTTCTTAAATTTCTCATCCCTCCAGGGCTCCCGTGCTATGTTCTCT
TTACCCCTTCCCCCTCTTCTCTTGCTCAGGCCTGCACCACTGCAGCCACCGTTTCATTTATTCA
TTCATTAAACACTGAGCACTCACTCTGTGCTGGGTCCCGGGAAGGGTGAGGGGGTCAGACACA
GGCCCTGCCCCCTGCCCTCAGTGACTGGCCAGTCCAGCCCAGGCGGGGAGAGATGTGTACATAG
GTTTTAAAGCAGACCCAGAGCTCATGGGGGCTGTGTTCTGGGTGTTCAAGGTGCTGCTGGTCC
TCCATTACCCACTGCTCCCCAAGGCTGGTGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTC
CTTCCCGTTCTCTCATCCACCTGCCCAGTGCTCATCGTTACAGCAAACCCAGGGGGCCTTGGC
CAGGTCAAGGGTTTCTGTGAGGAGAGGACCCAGGAGTGTGGGGGCATTGGGGGGGTGAAGTGGC
CCCCGAAGAAATGGAACCCACACCATAGCTCTCCCCACAGCTGATACGGCATCTTCGAGAAAG
ACCTGCCCTCTCACTGGGATCCCCCTTCTGCTCTCTCCAGGGCTCTGCCAGGGCCTTGCTC
AGTCCCTTCCACCAAAGTCATCTGAACCTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACA
CTGATGTCTGTCCCCAGGTGCTCTCTGCCCCCTCATGCCCCCTCTCACCGGCCAGTGCCCCGAC
TCTCCAGGCTTTATCAAGGTGCTAAGGCCCGGGTGGGAGCTCTCGTCTCAGAGCCCTCCTC
CGCCTGGTGCTGCCCTTTACAACACCTGCGAGGAGAAGGGCCACGGAAGCCAGGCTTTAGA
GCCCTCAGCAGGTCTGGGGAGCTAGAGCAAAAGGAGGACCTCAGGCCTTCCGTTTCTTCTTCC
AGGGTGGGGTGGCCTGGTGTTCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAG
AGGGAGGCGGCCTCCGCCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCGGGGGGT
GATCTCTGGTGCTCACAGCCGAGGGAGCCGTGCTCCATGGCCAGATGACGGAACAGGGTCT
GACCAAGTCCAGGAAGACCTGTGTATATAAACACCTGCCTGATCTGCCCTGCCTGACCC
CGCCACGCGCTGCCGTCCAGCATGATTAAAGAATGCTGTCTCTCTTGGAAAAA

62/615

FIGURE 62

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPHA
LPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRFFA
FSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKETVH
IMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYITFSGH
LIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 63

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGAC
 TCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCACCA
TGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCCTTG
 ATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTGTCAC
 TGCTGCCATTGAAGTTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGCACTGG
 CCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCTTTAAGT
 GTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGACGAGTACC
 GCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTTACAGCTGCTTCGTGGAAGA
 CCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCCAACTGGGTTTTCC
 CAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTCGGGGAGGAGT
 TTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCACTCAGTATATG
 TGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCAAGTGCACAGCCTGTGGTCATAGAA
 GGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAGTGGCCCTGGCAGG
 CCAGCCTTCAGTTCAGGGCTACCACCTGTGCGGGGGCTCTGTCATCACGCCCCCTGTGGATCA
 TCACTGCTGCACACTGTGTTTTATGACTTGTACCTCCCCAAGTCATGGACCATCCAGGTGGGTC
 TAGTTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAGATTGTCTACCACAGCA
 AGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCTGGCCGGGGCCACTCACGT
 TCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGAAGTTCCCCGATGGAAAAG
 TGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGACGCCTCCCCTGTCTCTGAACC
 ACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAGGGACGTGTACGGTGGCATCA
 TCTCCCCCTCCATGCTCTGCGCGGGGCTACCTGACGGGTGGCGTGGACAGCTGCCAGGGGGACA
 GCGGGGGGCCCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAGTTAGTGGGAGCGACCAGCTTTG
 GCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACACCCGTGTACCTCCTTCTGGAAGT
 GGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAAGAGGAAGGGGACAAGTAGCCACCT
 GAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCTGGACTCCCCTGTAGGAACCTGCACA
 CGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCAAGTAGCAGGCCCGAAAGAGGCACCCTT
 CCATCTGATTCCAGCACAACTTCAAGCTGCTTTTTTGTTTTTTTGTTTTTTTGAGGTGGAGTCT
 CGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAATCCCTGCTCACTGCAGCCTCCGCTTCCC
 TGGTTCAAGCGATTCTCTTGCTCAGCTTCCCCAGTAGCTGGGACCACAGGTGCCCGCCACCA
 CACCCAATAATTTTTTGTATTTTTTAGTAGAGACAGGGTTTTACCATGTTGGCCAGGCTGCTCT
 CAAACCCCTGACCTCAAATGATGTGCCTGCTTCAGCCTCCACAGTGTGGGATTACAGGCAT
 GGGCCACCACGCCTAGCCTCACGCTCCTTTCTGATCTTCACTAAGAACAAGAGAGCAGCAAC
 TTGCAAGGGCGGCCTTTCCCACTGGTCCATCTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCT
 GACGAGATAAGCAGTTATGTGACCTCACGTGCAAAGCCACCAACAGCCACTCAGAAAAGACGC
 ACCAGCCCAGAAGTGCAGAACTGCAGTCACTGCACGTTTTTCATCTCTAGGGACCAGAACCAAA
 CCCACCCCTTCTACTTCCAAGACTTATTTTACATGTGGGGAGGTTAATCTAGGAATGACTCG
 TTTAAGGCCTATTTTTCATGATTTCTTTGTAGCATTGTTGGTGTGCTTGACGTATTATTGTCCTTTGA
 TTCCAAATAATATGTTTCCCTCCTTCAATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAAT
 CAAAATCATCCACTGAAA

64/615

FIGURE 64

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234
><subunit 1 of 1, 453 aa, 1 stop
><MW: 49334, pI: 6.32, NX(S/T): 1
MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILAL
AIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLOVFTAASWK
TMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKVTALHHSVY
VREGCASGHVVTLOCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVITPLWI
ITAAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKLAGPLT
FNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPLISNKICNHRDVYGGI
ISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSFQIGCAEVNKPQVYTRVTSFLD
WIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

65/615

FIGURE 65

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGT
GAGGTTGGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAG
AACGCTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCC
ACACAACTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC
AGGAGGACTTTCTGTTTGTGTTGTCACCTTTGACCTCTTATTTCGTAACATTACTGTGGATAATA
GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC
TATTCTTCATATTTTGATATATTTCTTCTGTCAGTTTTTTCGATTTAAAGTGTTAATACTTGCA
TATGCTGTGTGCAGACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC
TTTTTACTAGCAAAAGTGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTG
CCCATCATTTTCATTCATCCTTGCCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCT
CAAGAAGCAGAAGAAGAAAAACAGACTCCTGATAGTTTCAGGATGCTTCAGAGAGGGCAGCACTT
ATACCTGGTGGTCTTTCTGATGGTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAA
GAAGCTGAAGAAAAACAGGACAGTGAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTA
AATGTGAAAAACCCCTCACAGAAAGTCATCGAGGCAAAAAGAGGCAGGCAGTGAGTCTCCCTG
TCGACAGTAAAGTTGAAATGGTGACGTCCACTGCTGGCTTTATTGAACAGCTAATAAAGATTT
ATTTATTGTAATACCTCACAAACGTTGTACCATATCCATGCACATTTAGTTGCCTGCCTGTGG
CTGGTAAGGTAATGTCATGATTCATCCTCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAAA
TAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATCAAAAGACTTAATATATTGAAGTAACACTT
TTTTAGTAAGCAAGATACCTTTTTTATTTCAATTACAGAAATGGAATTTTTTTTGTTTCATGTCT
CAGATTTATTTTGATTTCTTTTTTAACACTCTACATTTCCCTTGTTTTTAACTCATGCACA
TGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTGACATGTCAATGTGGCTAGTTTTA
TTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGACTTGCAAAAGGGGAAGAAAGG
AATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTTTTATCATGAAATCATGTTT
TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGCACAAAATGACTTAAACC
ATTCATATCATGTTTCCTTTGCGTTTCAGCCAATTTCAATTAATAATGAACTAAATTAATAA

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGT

66/615

FIGURE 66

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDL
LFVTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAI
ALTTAVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIV
QDASERAALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 67

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCATC
CAACCATTCCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCACCCGACT
AACATCTCAGTCTCTGAAAATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCAC
GGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTACCTGC
CACCCCTCAACGTCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGCTACAC
AGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGA
GATGTTCCCTCCAGTTCGCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGT
GGAGTTCTCAGGGAACCCCAGCAAGTACGATGTGTGCGGTGATGCTGAGAAACGTGCAGCCGGA
GGATGAGGGGATTTACAACCTGCTACATCATGAACCCCCCTGACCGCCACCGTGGCCATGGCAA
GATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGGCCGTGATTGT
GGGTGCCTCCGTGCGGGGCTTCTGCGCTGTGGTCACTTTGGTGCTGATGGTGGTCAAGTGTGT
GAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGAGGAGGGCAAGAC
GGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCCCTGCAGCCTCCCG
TGTCCCGTCTCCTCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCGCTCTTGGTGTGCTT
CCCGTGACCTAGGACCCAGGGCCACCTGGGGCTCCTGAACCCCCGACTTCGTATCTCCCA
CCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCATGCTCTGGGACGTGTG
GGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGGGGGGAGGCAGGAGGCAC
ATGTGAGGGTCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGAGGAGGGGCCGTGTACC
TGCCAGTGCTTGCCCTGGCAGTGGCTTCAGAGAGGACCTGGTGGGGAGGGAGGGCTTTCCTGT
GCTGACAGCGCTCCCTCAGGAGGGCTTGGCCTGGCACGGCTGTGCTCCTCCCTGCTCCCAG
CCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGAACTTGGAGGGGCATGTTAAA
GGGATGACTGTGCATTCCAGGGCACTGACGGAAGCCAGGGCTGCAGGCAAAGCTGGACATGT
GCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATTGCTAGTGGCCTCCTTGGGGCTCC
TGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAGACTGGAAGAGCAGCTCCAGGTAGG
GGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCCAGTTTCAAAGTCAGCTGAGGGGCTG
AGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAGGCTCTGCCTTCTCCATGGGGTAACCA
CCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAATGAGGAGGCCATGCACAGGGTGGGGCAG
CTTTCTTTGGGGCTTCAGTGAGAACTCTCCAGTTGCCCTTGGTGGGGTTTCCACCTGGCTTT
TGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCGGCATAAGGGGAGGCCTTGGAACCTGAGCT
GCCAATGCCAGCCCTGTCCCATCTGCGGCCACGCTACTCGCTCCTCTCCCAACAACCTCCCTT
GTGGGGACAAAAGTGACAATTGTAGGCCAGGCACAGTGGCTCACGCCTGTAATCCCAGCACTT
TGGGAGGCCAAGGCGGTGGATTACCTCCATCTGTTTAGTAGAAATGGGCAAAACCCCATCTC
TACTAAAAATACAAGAATTAGCTGGGCGTGGTGGCGTGTGCCTGTAATCCCAGCTATTTGGGA
GGCTGAGGCAGGAGAATCGCTTGAGCCCGGAAGCAGAGGTTGCAGTGAACTGAGATAGTGAT
AGTGCCACTGCAATTGAGCCTGGGTGACATAGAGAGACTCCATCTCAAAAAAA

68/615

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415
<subunit 1 of 1, 215 aa, 1 stop
<MW: 24326, pI: 6.32, NX(S/T): 4
MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQF
SLNWTYQECNNCSEEMFLQFRMKI INLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDEGIYN
CYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRKKEQ
KLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

10055556 014506

FIGURE 69

GGCGCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTGCGCGCCAGCCTCCGCCGCCGAGCCTC
 GTTCGTGTCCCCGCCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCCCACCCTGGCAGACTAACGAA
 GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC
 CGCAGAGGCGGAGGCTCGCGTATTCTTGCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
 GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGCTCCGGTTTGGCT
 CACCTCTCCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
 ATCCTGAGGTCATTCATTATGAAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGTCTTTCATGGCTAGA
 GCAATTCAGCCATGGTGGTTCCCAATGCCACTTTATTTGGAGAAAATTTTGGAAAAATACATGGATGAGGATGGT
 GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
 AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
 TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
 TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA
 CAGGTGCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCTATAACATGAACATCTGGGGGCGAG
 ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAACTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGCCCTGTTCTGCTTGGCCACCTAGTTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
 TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
 CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTTGTTTCT
 TGTGAAGTAAGATTAAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAATGTCTGCTGGCTGTTTGGAT
 AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
 ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAAATAGA
 AATGGTATTCAAACAATTTGGCAAATATCAGTCTGCTAATTCCTTACAGTCTCTAAAGTAACAGTTTCAGGCTGTG
 ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTCATAAGCCTGCTTCACATTGCCCCAAGAGTATACTGTCCT
 CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATTTGGAACCTCGAGTTTATTTCTGATCTGTCCAGTATC
 TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCCTGTGGACAAAAGA
 AAGACCTACATTGCTTCTTTTTCAGAAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
 AGAGTGTGTTGCTGTTGTGTGAAGTGAATACTTGGAAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTAAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCCAAAAGGTGCCAAATGCATA
 TAAATCTTGATAAACAAAGTCTATAAAAATAAAACATGGGACATTAGCTTTTGGGAAAAGTAATGAAAATATAATGG
 TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
 GTTCTACGTTTTCATATATTATATGGTGTCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
 GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACCTGAAAGAAACCTTATCACATTTTCCCC
 AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
 TGTTAATTTTAGGCATATAGAATATTAAATCTGATATTGCACTTCTTATTTTATATAAAAATAATCCTTTAATATC
 CAAATGAATCTGTTAAATGTTTGTATCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
 ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
 AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
 GAACAAAGATGAACTAATGTATTACATTACCATTGCCACTGATTTTTTTTTTAAATGGTAAATGACCTTGTATATAA
 ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
 AATCTGTAAAATGTTAGTTTTTGGTAATTTTTTTTCTGCTGGTGGATTACATATTAAATTTTTTCTGCTGGTGA
 TAAACATTAAAAATTAATCATGTTTCAAAAAAAAAAAAA

THE OF THE **WORLD** AND ITS PEOPLE

FIGURE 70

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLEKYMDEDEGEWWIAKQRGKRAITDNDMQ
SILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPA SLLPSIGQNLGAHWGR
YRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLCHNM
NIWGQIWPKAVYLVCNYS PKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRYPPR
EEETNEIERQQSQVHDTHVRTRSD DSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNRYECPA
GCLDSKAKVIGSVHYEQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGIQTIGKY
QSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIGTRVYSDL
SSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFONGIFSES LONPPGGKA FRVFAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

CAGCCCCGCGCGCCGCGCGAGATCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGGC
TGGGCGCGCCCCCGGGCCCCGCGGTGGGC**AT**GGGCGCACTGGCCCCGGGCGCTGCTGCTGCC
CTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGAGCTGGCCCCCGCGCCCTTCACGCTGCC
CTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCGGGACCCGGGACCCCTGCC
GAGCGCCACGCCGACGGCTTGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCC
GCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAG
ATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACCTTT
GCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACA
TACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTT
GGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT
ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT
TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAA
GCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGCTTGCCCGTTGCTGGATCT
GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATC
TGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAATTGGA
GGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGC
ACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAAGCTGTGGCCCGCGCA
TCTCTGATTCCAGAATTCTCTGATGGTTTTCTGGACTGGGTCCAGCTGGCGTGCTGGACGAAT
TCGGAAACACCTTGCTTACTTCCCTAAAAATCTCCATCTACCTGAGAGACGAGAACTCCAGC
AGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACGCCATGATGGGGGCCCGCCTG
AATTATGAATGTTACCGATTTCGGCATTTCCTCATCCACAAATGCGCTGGTGATCGGTGCCACG
GTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGACGCGAGC
CCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCCTTCTCAACAGAGGAT
GTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTAT
GCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTT
CGGTGTCAGCGTCGCCCCCGTGACCCTGAGGTGTCATGATGAGTCCTCTCTGGTCAGACAT
CGCTGGAAT**TGA**ATAGCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCA
CATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCACCCGTCTTCAAT
CTCTGTTCTGCTCCCAGATGCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTTCAAGCTTT
CAAATCCTCCTACTTCCAAGAAAAATAATTAATAAAAAAAAAACTTCATTCTAA

72/615

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLALA
LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYI
DTYFDTERSSTYRSKGFDTVVKYTQGSWTGTFVGEDLVTIPKGFNTSFLVNIATIFESENFLLP
GIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLG
GIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLLRLPQKV
FDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISYLRDENSSRSFRITILPQ
LYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFASPCAIEIAGAAV
SEISGPFSTEDVASNCVPAQSLSEPIWIVSYALMSVCGAILLVLIIVLLLLLPFRCQRRPRDPE
VVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CCGCGCGGCGAGAGCGCGCCAGACCCCCCGCCGCGATGCCCCGCGCGCCAGGACGCTCCTCCCCGCTGCTGGCCCCGGC
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGGCGCTGGGGCGCCCGGGCCCAGG
AGGCGGCGGCGGCGGCGGCGGAGCGGGCCCCCGCGGCAGACGGCGAGGACGGACAGGACCCGCGACAGCAAGCACCC
TGTACACGGCCGACATGTTTACGACACGGGATCCAGAGCGCCGCGCACTTCGTATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGTGCGAGCCGACTTGAATGACCTGGGAGACAAATCAACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGAGCTGTGCTCCGCCCCAGGGGGTGCGAGGATACCCCACTTTAA
AGCTTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCTTCGGGACTTCCAGACACTGGAAAACTGGATG
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGAACCGCCCACTGCCCCCGAGCTCAAGCAA
GGCTGTATGAGCTCTCAGCAAGCAACTTTTGAAGTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC
CTGTGTGTGGTCACTGCAAGGCCCTGGCTCCAACCTGGGAGCACTGGCTCTGGGCCTTGAACATTTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTATACACAGCACTATGAACCTGCTGCTCGGAAACCAAGGTTGCTGGCTATCCCCACT
TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGGAAAGCGGGATTGGAGTCACTGAGGGAGTACG
TGGAGTCGCAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCAAGCCCCAGAGGGCCCCGGTGTGCGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAAAGTTCGATGACACCATTGCAGAAGGAA
TAACCTTCAATCAAGTTTATGCTCCATGGTGTGGTCAATTGTAAGACTCTGGCTCCTACTTTGGGAGGAACTCTCTA
AAAAGGAATCTCCCTGGTCTGCGGGGGTCAAGATCGCCGAGTAGACTGCATGCTGAACGGAATATCTGCAGCA
AGTATTCGGTACGAGGCTACCCCACTGTTATGCTTTTCCGAGGAGGGAAGAAAGTCACTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGTCTTGAGCCAAGCGAAAGACGAACCTTTAGGAACACAGTTGGAGGTCAC
CTCTCCTGCCAGCTCCCGCACCCCTGCGTTTAGGAGTTAGTCCCACAGAGGCCACTGGGTTCCAGTGGTGGCT
GTTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTTCCAAGCCAACACACTCTACAG
ATTTCTTTATTAAGTTTAAAGTTTCTCTAAGATAAATGTGTAACCTCATGGTCACTGTGTAAAACATTTTCAGTGGCGATA
TATCCCTTTTGACCTTCTCTTGATGAAATTTACATGGTTTCTTTTGAGACTAAAATAGCGTTGAGGAAATGAAA
TTGCTGGACTATTTGTGGCTCTGAGTTGAGTGATTTTGGTGAAGAAAGACATCCAAAGCAGGTTTACCTGCG
CCACGAGTTCTGGAAGGTTGGCCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGCGCGCTGC
TGGAGCTTCTGTTGCTGTAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCCCTGAGTAAGAGAAATGCTGATGCCATAACTTTATGTGTGATGCTTGTCAAAATCAGTTACTGTTTCAGGGGAT
CCTTCTGTTTCTCACGGGGTGAACATGCTTTTAGTTCTCATGTTTAACACGAAGCCAGACCCACATGAACTGT
TGGATGTC'TTCTTAGAAAGGGTAGGCATGGAAAATTCACAGGCTCATTCTCAGTATCTCATTTAACTCATTGA
AAGATTCCAGTTGTATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTTCCAGGCCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAAGTAGAGTTCTAGAGTTTCTGATTCGTGTTTCTCAGTAGTCCTTTTAGAGG
CTTGCTATACCTTGGTCTGCTCAAGGAGGTCGACCTTCTAATGTATGAAGATGGGATGCATTTGATCTCAAGAC
CAAGACAGATGTCAGTGGGCTGCTTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTGTATGCCAGTGTCTCTA
ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCCATAATAACCACTTTGCATCCAACACTCTTACCCACCT
CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCAG
TATGGTTTCAAGATAATTTCTTTTTTAAAAAAACCCCAACCTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTACA
CACAACTTTCAGCTTTGTCATCAGAGTCTGTGATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTTAA

74/615

FIGURE 74

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0
MPARPGRLLPLLARPAALTALLLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHSK
HLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAHS
VCSAQGVIRGYPTLKLKFKPGQEAVKYQGPRDFQTLNWMQLTLNEEPVTPEPEVEPPSAPELKQ
GLYELASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCTQHYE
LCSGNQVRGYPTLLWFRDGGKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEAPVLAA
EPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLAGVKIAE
VDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL
```

Signal sequence:
amino acids 1-32

20250101 09:00:00

CGGACGCGTGGGCGGACGCGTGGGCAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGGCTTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCTCTCACTTCGCCGCGCTGGTGAGTGTTCGGGGAGATTGGCAAACGCCTAGG
AAAGGACTTGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGCTCCACTACGGCAGTTTATCTC
TCTGATCAGAGCGACAGCGCGCTCCACTTCGCAGTTCCTTCCAGGTGTGGGACCGCAGGACAGACGCCCGGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCTCTGTCCACTGTGATTCTCAATTCTTGTCTTGGTTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTTCTTTTTTTTTTTTTTCTTGGTGGAAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTCTCTCCGAGTCCCATCCATTAAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTCTTGGCCAGGTGGAGAGCTTCCACCGTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGCGCGGGCGGTGGTTCTCCGCTGGAGTCTCAGCTGGGACCTGAGTGAATGCTGCCATCCCA
GGGGCTGTGCGGGGCATCCGCTCCGCTTCTCCACAGGCCGTGTGTCTGTCTGGAAAGATGCTAAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTCTGGGGCCAGGCCTTAGAAGAGGAG
AAGAAGGGGCCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA
TTTTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGGTTACCACGGATCTGAGATTA AAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAACTGGAGAACTACTATGTCCAGCCTATTTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTCTATCATAGACCTACCCAACTCCAACTGTTTAC
CTCTGGACAATGCCACCTTACCTCAGAAACTGAAGGAGTTGGATATTCAACGCATATGGTCGGAAAATGGCACT
TGGGTTTTTAAACAGAAAAGAATGCATGCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACACTACAAATGTGACAGTCTCTGGGATGTGTGGCTATGACTTGTATGAAAACGCAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCAT
ACCCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTTCACTACCACCTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAAACAGGAGAAGATATGCTGCCATGCTTTCTGTCTTAGATGAAGCAA
TCAACACAGTGCAGATTGGCTCTAAAGACTTTATGGTTTTCTATAACACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAAGTGGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCATG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTTCAGCAACCTGGGACCGAACC
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCAGAGTTCAACAAAC
TGCAGTGCCGGTCAGGTATCCCCCAAAGACCCAGAAGTAACCTTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAGCCAAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAACCCAGCAAAATTTGGCTCGATAATATCGCTGGCCTAGCGCTCA
GGCTTGTTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACTTGGCCGCCACACTGAAAACCTGTCTCTGTCAGT
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCTTGAGTTCCACTGCTGTGCTTCAGTCAACTGACCAAACACTGCTTTGAATTATAGGAGGAGAAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTA AAAACTACCTTTGATAAATTAC
AGTCAAAGATTTGTGTCACTCCAAAGGCCCTGAAGAATATATTTTCTTGGTGAATTTTGTATGTCTGTATATGA
CACTTGGGTTTTTTAATTAATCTATTTTATATATATAAATATATGTTTTCTTTCTGTGAAAAGCTGTTTTTCT
CAGATGTGAACAGCTTGCACCTCATTTTACCATCGCTGAGGGAATGGCAAATAAGAATTGTTTGAACACACTGCC
ACAATGAATGTAACATATTTTCTAAACACTTTACTAGAAGAACATTTTCACTATAAAAAACCTAATTTATTTTACA
GAAAAATATTTTGTGTGTTTTTATAAAAAGTTATGCAAATGACTTTTATTTTTATTTTCTGTCATACCATTAGAAGA
ATTTTATTTTATTTCTTCAAATTATCAAGCACTGTAATACTATAAATTAATGTAATACTGTGTGAATTCAGACTA
TAAAAAACATCATTCAGAAAACCTTTATAATCGTCAATGTTCAATCAAGATTCTTTGAATGTAATAAGATGAATATAT
ATTACTTGGAAATTCATGTGTTTGTGCAGAGTGTGAGCAAACTTTATGTGTTTCTATCATAAACTATTATGATCTT
AATTTATTAATAATGATTTACTTTTATGGCACTAGAAAATTTACTGTGGCTTTTCTGATCTAACTTCTAGCTAAAATT
GTATCATTTGGTCCTA AAAAATAAAAAATCTTTACTAATAGGCAATTGAAGGAATGGTTTGTCTAACAAACCACAGTAA
TATAATATGATTTTACAGATAGATGCTTCCCCCTTGGCTATGACATGGAGAAAGATTTTCCCATATAATAACTAA
TATTTATATTAGTTGGTGC AAAAATAAGTTGCGGTTTTTCCCATTA AAAGTAATAACCTTACTCTTATACAAAGT
GGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCTCTGCTGGAGTAGGTAACTTGTCTTGGAAACCCC
ACATGCAAAACGTCATGAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTTATCATGGGTCATCAATGAGCATA
GATTGGTGTGGATCTCTGTAGACCTGGTGTCTTTTGAAGTGCCTCTCCTAATGACAGAGGCCTTGAAGCTTAC

76 / 615

FIGURE 75B

[illegible][illegible]

77/615

FIGURE 76

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEEGALLAQAGEKLEPSTTSTSQ
PHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAAGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQP
NCLPLDNATLPQKLKEVGYSTHVMVGKWHLGFNKECMPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCYDLYEN
DNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSC
DEAINNVTLALKTYGFYNNSIIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCHEL
HITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPGQQAMGSGTLQSSQPSE
CSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFSTSQPTHMRGWYTLTG IQES
```

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

20250909 10:55:00

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCATCCCTTTTGAAGAACAGTACTGTGGA
 GCTATTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
 GCCGCTGTTACCAATCGGGGAGAGAAAAGCGGAGATCCTGCCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGAT
 AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT
 GGGCGAAGGAATGCTCTGACTTCACTGTTAGTGGTAAAGGCGAGAATGAAAAATTAATCTGGAGGGAAGATAAGAATGAT
 TCTCTGCGCAGTGCACCGGACTACAAAGGGCTTGTCTGTCTGGGAATCCTCTGGGAGTCTCTGTGGGAGACCGG
 ATGCACCCAGATACGCTATTAGTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGACCT
 GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCGAGCTTTTCGCCC
 GAATCCGCGCAGCGGCAGCTTGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
 TCAATTAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAAGTAGAAGTAGAAGTAAGGGACATTAA
 CGACAATGCGCCTTACTTTTCGTGAAAGTGAATTAGAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT
 CCTCTACCCCAAGCTGGGATATCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGCAACACTCA
 CTTCTCCTCATCGTGCAAAATGGAGCCGACGGTAGTAGTAAGTACCCCGAATTTGGTGCTGAAACGCGCCCTGGACCG
 CGAAGAAAAGGCTGCTCACCACCTGGTCCTTACGGCCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGCG
 CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCAGCGTTTGCTCAGCCCGAGTACC GCGCGAGCGTTCC
 GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCCTGACGAAGGAGTCAATGCGGAAGT
 GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCAAGTTTTCAAAGTATGTAATTCAGGGACAATATC
 AACATAGGGGAGTTGGACCAAGGAGTCAAGGATTTACAGATGGAAGTGCAAGCAATGGATAATGCAGGATA
 TTCTGCGCGAGCCAAAGTCTGATCTGTTCTGGACGTGAACGACAATGCCCAGAAGTGGTGCTCACCCTCTCT
 CGCCAGCTCGGTTCCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAATGACCAAGATTCTGA
 GGAAAACGGACAGGTGATCTGTTTTCATCCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA
 TAGTTTGTAGTACAGACATAGTCTTGGATAGGGAAACAGGTTCTTAGCTACAACATCACAGTGACCGCCACTGACCG
 GGGAACCCCGCCCCCTATCCACGGAACCTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCCGGTCTT
 CCTCAGGCTCCTTATTCGCTTATATCCAGAGAACAAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
 CGACCCCGACTGTGAAGAGAACGCCAGATCACTTATCTCTGGCTGAGAACACCATCCAAGGGGCAAGCCCTATC
 GTCTACGTGTCATCACTCCGACCTAGGCTGTTTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
 CTTGCAAGTGAAAGTGATGGCGCGGGAACAACGGGCACCCGCCCTCAGCAGCAACGTGTCTGTTGAGCCTGTTCTGT
 GCTGGACGACAACGACAATGCGCCCCGAGATCCTGTATCCCGGCCCTCCCCACGGACGGTTCCACTGGCGTGGAGCT
 GGCTCCCCGCTCCGCGAGAGCCCGGTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
 CTGGCTGTCTTACCCTCTGCTCAAGGCCAGCGAGCGCGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
 CACGGCGCGAGCCCTGCTGGACAGAGACCGGCTCAAGCAGAGCCCTCGTAGTGGCCGTCAGGACCAAGCCGACGC
 CCCCTCTCCGCCACTGTACGCTCACCCTGGCCGTGGCCGACAGCATCCCCAAGTCTCTGGCGGACCTCGGCAG
 CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG
 CGTCTTCTTGGCCTTCTGTATCTTGTCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTACGCCTGCTGCAGGC
 TTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTGCACATTTGTGGGCGTGGACGGGGTGCAGGCTTTTCTGACAGC
 CTATTTCCACAGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACTTGATCTTCCCCAGCCCCAACTATGCAGA
 CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAGCGAGCCCCCTTTGTCTGTTCAGGTGATTTCGGTATTTTCTAAAGA
 CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTTTTTTAATTGCTCTGTCTCCCAAGC
 TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAACCTCCTAGGCTCAAGCAATTATCCACCTTTTCCCT
 CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCATATCTATCTATCTATCTATCTATCTATCTAT
 CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCAGCCTGTAATCCAGTACTTTGGGAGGC
 CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAGACCAGCCTGACCAACATGGAGAAACCCCGTCTATACTAA
 AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCAGCTACTTGGGAGGCTGAGTCAGGAGAAAT
 TGCTTTAACTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACCTCCAGCCTGGGCAACAAGAGTG
 AAACCTCATCTCA

79/615

FIGURE 78

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPPEELEKGSRVGDISRDLGLEPRELAERGVRRIIPRGRTQLF
ALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVRDINDNAPYFRESELEIKISENAATEM
RFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGADGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGT
ARIRVMVLDANDNAPAFAPQPEYRASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGT
ISTIGELDHEESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVDQD
SEENGQVICFIQGNLPPFKLEKSYGNYISLVTDIVLDREQVPSYNITVTATDRGTPPLSTETHISLNVADTNDNPP
VFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTIQQASLSSYVSINSDTGVLYALSSFDYEQF
RDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDNAPEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQ
NAWLSYRLLKASEPGLFSVGLHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADL
GSLESPANSETSDLTLYLVVAVAAVSCVFLAFVILLALLRLRRWHKSRLQASGGGLTGAPASHFVGVDGVQAF
QTYSHVSLTSTDRLSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGLIEVSLYQIFFLFFFNCSVS
QAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDRCRKPPTVCLSIYLSIYLSIYLSIYLLLSCTDGSGLTPVIPVLW
EAEAGGSPEVGSRLPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 79

AGCCGCTGCCCCGGGGCGGGCGCCCCGCGGCGGCACCA**ATG**AGTCCCCGCTCGTGCCTGCGTTTCGCTGCGCCTCCTC
GTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCAAGCTGTCTGTCGGTGGGGAGCATCTCA
GAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATCCAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTC
ATGGACTCGGTGCGCCGCGGTGCCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAC
TGCTCCACACTCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCCTTCGTGTAC
GCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCTGGAGAAGTGCGGCTGT
GACAGGACAGTGATGGGGTGAGCCACAGGGCTTCCAGTGGTCAGGATGCTCTGACAACATCGCCTACGGTGTG
GCCTTCTCACAGTCGTTTGTGGATGTGCGGGAGAGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTC
CACAACAATGAGGCCGGCAGGAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTGAGGC
TCCTGTGAGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAGAAGTTT
GATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACCACGCAACGCACAGTTCAAG
CCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCGACTTCTGTGAGCAGGACATGCGCAGCGGCGTG
CTGGGCACGAGGGGCCGCACATGCAACAAGACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGC
GGCTTCCACACGGCGCAGGTGGAGCTGGCTGAACGCTGCAGCTGCAAAATTCACCTGGTGTGCTTCGTCAAGTGC
CGGCAGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCG**ATGA**CCGCCCTGCCTAGCCCTGCGCCGGCAACCACTA
GTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGAGATACTGGTTGTATTTTTTTGTT
CTGGTTTGGTTTTTGGGTCTCATGTTATTTATTGCCGAAACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGC
CTCCCCAAAGCCTGGGCCTTTGTGGCTGCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTG
GCTGCCACTGACCACCTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGACAGCTAAGGTGGAGTAACAAGGAGTA
TTACCACCACATGGCTACTGACCGTGTCTCGGGGAAGAGGGGGCCTTATGGCAGGAAAAATAGGTACCGACTTG
ATGGAAGTCACACCCCTCTGGAAGAAAGAACTCTTAACCTCTCCAGCACACATACATGGACTCCTGGCAGCTTGA
GCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGGGAACAAGCAGATACCAGGTCAAGGGCACCAAGGTTTATT
TCAGCCCTTACATGGACAGCTAGAGGTTCGATATCTGTGGGTCTTCCAGGCAAGAGAGGGAGATGAGAGCAAG
AGACGACTGAAGTCCCACCCCTAGAAGCCAGCCTGCCCCAGCCTGCCCCCTGGGAAGAGGAAACTTAACCACTCCCC
AGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTTTGCAGTCATGCCCCGA
GTCACCTTTTACAGCGCTGTTCCCTCCATGAAACTGAAAAACACACACACACACACACACACACACACACACAC
ACACACACACGGACACACACACACACCTGCGAGAGAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCGAG
TCACCTTTTACAGCACTGTTCTCTC

81/615

FIGURE 80

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLEVMDSVRRGAQLAI
EECQYQFRNRRWNCSTLDSLVPFGKVVTQGTREAAFVYAISSAGVAFVTRACSSGELEKCGCDRTVHGVSPQGF
QWSGCSDNIAYGVAFSQSFDVREERSKGASSSRALMNLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTWCWRAVP
PFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTS
KAIDGCELLCCGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

82/615

FIGURE 81

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCAATCAAGAAT
TGGGGTCTGGCTCAGAAATCTCGCAGCTGGTGAAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTTATGAAGCTCTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTATGCCCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTTATTCGCCACGGAGACAGGTACCCACTGTATGTCAATCCCA
AAACAAAGCGACCAGAAATGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTCA
TTAGTCACATGTCAAAGGATCCGGAGCCTCTTTCGAAAAGCCCCCTGAACTCCTTGCCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCCTGGGAAAA
GCCGGACCCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCCTGTTCTGCTCTGGAAGCTGCTATTGCCCGTAAGAAACAGTATCTGGAAGGAGC
AGCGTCGTGAGTACCTCCTACGTTTGAAAAACAGCCAGCTGGAGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTTGACTTCGGGTATTCTCTCCTGGGTGCCACCCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTGGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTGGGCCTTTGAGAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCCAGTGAACATTCGGTCCGGATTCTTTACAATGGCGTCGATGTCAATTTCCACACCT
CTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACCTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTACAGGGAAGGATTCTAAAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATGCTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAATTTTTTTAATCTTAGACATT
TTTACCTTGTCTTGTAAAGAAATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTTCACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTACATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTCCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCCCTGGGCTGTAGTTTGTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCCTCCAGTAACCTCTGCTAGAAAACA
CAGAAATTTGGTCTGTATCTGACACTAGAACAAAACTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTtaggaaaatatt
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACCTTTCTCCAATAAACTAATTATGGCTCATTCCTTTGACAAGCTGTAGA
ACTGGATTCAATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTTCTGATTGATTTTTAAATGCGTTTTTGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACATAATGGTGCTAATTCAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTATCGGCATTCC
AACTTTTTCTCTTTGTTTTGTCCAGTGTGCAATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAGAATAA

1005556 011502

[illegible]

85/615

FIGURE 84

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGMEEQLNKIPGFCENE
KGVVPCNILVGKYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNGFWFFKFAAAIAIIIGAFFIPEGTFTT
VWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKMEEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPAS
CSENKAFISVNMLLCVGASVMSILPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSLLSIIGYNTT
STVPKEGQSVQWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVHRAV
DNERDGVITYSYFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSWIGIVLYVWTLVAPLVLTNRDFD

10052336-04430

AACAAAGTTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGAGCTGCTGGGCAGAGAGGGGACTTCCCGCTCCCAGATGCTGGGCCCTCCTGGGGAGCAGACGCCCTCTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCGCCACCTGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGGAAACCGAGTCCGCCGGGCCAGCCTTTGGCCCTTCCGGCGGGCGGGGCCACTTGGGAATCTTTTACCATCACCGTCACTCTGGCCACGTATCTCATGTGCCGAATGTGGGCCCTCCACCACCACCACCACCCCGCCACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCGGCCACGCTCGCTGAGAGGCTGCTGTGCGCGGTGCTGTGGACAGCAGCTGCCCTTGCCCTCCCATCTGTGTCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGTGAACGAGGGGAAACAATAGACTGGGCTTGTCCAGCTGCATTTGATGGCATGCCATGCCCCAGTGTACTATGGCAGCAGAGAAATGGAGGAACACTGGGTCTGCAGTGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGGGTGCTCTTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAAGTGAGCCCCAGAAATGACAAGCGTGTCTTGGCAGAGCCAGCACAAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTCATCAGGCTGTGCTCAGGCCCTCTGGCGGGCAGGCGCACTGGGAGAGGCCCTGAGAAATGTCTTTTGGTTTGGAGGAAGGCAGTGTGAGGCTGCACAGTCAATTTCATCGGTGCCTTAGTCCAAGAAAAATAAAAAACCAATAAGAAGCTTTAAAAA

87/615

FIGURE 86

MLGLLGSTALVGWITGAAVAVLLLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAPWPFRRRGHLGIFHHHRH
PGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPHRHHPRHAR

1005555.0455

CCACCGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTCAAGGCTGGCTTGTGCCAGAACGGCGCGCGCGCGCACGCACACACACGGGGGAAACTTTTTTTAAAAATGAAAGGCTAGAAAGCTCAGCGCGCGCGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGACGCCCGGCCCGGCTCGGCGCCCCGCTGGGATGGTGACAGCGCTCGCCCGCGGGCCGAGAGCTGCTGCACCTGAAGGCCGGCGACAGTGGCAGCGCGCCGCTGCCCCGTGCCCCGCCCGCCCTCTGCTCGCCCTGGCCGGTGCCTCTGCTCGCCCTGCGAGGGCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGAGTAAGTTGTCAGTGCCCTCTGTTGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTTCGACTCCAAGAATCATCCAGAAGTGCTGAATATTCGACTACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTTCACGGAATCCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGTACGGGGATATTCTGATTCAAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAATGAAGCTATGTCTTTAGAACCAATGAAAAGTGCAACCAACAGATACAAACTCTTCCAGCGGAAGAAGCTGAAAAGCGTCCGGGGATCATGTGGATCACATCAACACCAACCACTCGCTGCAAAAGAATTGTGTTTCCACACCCCTCTCAGACTGCGGCAAGAAGTCTAAAAAGAGAGACCTCAAGGCACTAAGTATGTGAGCTGTGTGATCGTGGCAGACAACCGAGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTTGACAA GTTTTACAGACCACTGAACATTTCGGATCGTGTGGTAGGCGTGGAAGTGTTGAATGACATGGACAAATGCTCTGT AAGTCAGGACCCATTACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA TGACAAATGCGCAGCTTGTCAGTGGGGTTTATTTCCAAGGGACCACATCGGCATGGCCCCAATCATGAGCATGTG CACGGCAGACCACTCTGGGGGAATTGTATGACCACTCAGACAATCCCTTGGTGACGCCGTGACCCCTGGCCACA TGAGCTGGGCGCAAAITTTGGGATGTAATCATGACACACTGGACAGGGGGCTGTAGCTGTCAAATGGCCGTTGAGAA AGGAGGCTGTCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTAACTGCCGGAAGTCAGGGAGTCTTTTCGGGGGCCA GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG CTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT GAAGCCTGCGAGGAACAGCGTGCAGGGACTCCAGCAACTCTGTGACATCTCCAGAGTTCTGCAGAGGGGCCAGCCC TCACTGCCAGGCAATGTGTACCTGCACGATGGGCACCTCATGTAGGATGTGACAGCGTACTGCTACAAATGGCAT CTGCCAGACTCAGCAGCAGCAGTGTGTACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA GAGAGTCAATTTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTTCTTTGCCAAATGCGAGAT GAGAGATGCTAAATGTGGAATAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTC CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGGATTCTGTGCCGGGGGACCCACGTGTACTTTGGGCGATGA CATGCCGGACCCAGGGCTTGTGCTTCGAGGCACAAGTGTGCAGATGGAAATACTGCCCTGAATCGTCAATGTCA AAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGGCACGGCAGAGGGGTGTGCAACAACAGGAAGAA CTGCCACTGCGAGGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGCCCC CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTG CTGCTGCAGAGGAGGTACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT TCCATGACAAACAGACACAACACAGTTCTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCACGCTCTGCAGAAA CAGTGCAAGGAAGGGCAGCACTCTCTGGTTGAGCTTCTGCTTAAACACTGGACATGCTTCACTGCTGCTCTCTGAG AGAGTAGCAGGTTTACCACCTCTGGCAGGCCCGCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAGTCTGGCCTTTC ACTGAGCCTCCACAGCAGTGGGGGAGAAAGCAAGGGTTGGGGCCAGTGTCCCCTTTCCCAAGTGACACCTCAGCCT TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATTATATGAAAAT AGCAGGGTTTTTAGTTTTTAAATTTATCAGAGACCCCTGCCACCCATTCCATCTCCATCCAAGCAAACCTGAATGGCAA TGAAACAAACCTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGAACAGC AGTACTCAAGTTTGGAGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCCCTTCAATTAACAAGTAAGAA TGTTTAAAAAGTGAAACAATGTAGAGCCTAACTCCATCCCCCTGGCCATTACTGCATAAAATAGAGTGCATTT GAAAT

89/615

FIGURE 88

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624
><subunit 1 of 1, 735 aa, 1 stop
><MW: 80177, pI: 7.08, NX(S/T): 5
MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVVSASVRSGLWIPVKSFDKSKNHPEVLNIRLQRE
SKELIINLERNEGLIASSFTETHYLQDGTVDVSLARNYTGHCYHGHVIRGYSDSAVSLSTCSGLRGLIVFENESYV
LEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLAAKNVFPSPQTWARRHKRETLKATKYVELVIVADNREFQ
RQGDLEKVKQRLIEIANHVDKFYRPLNIRIVLVGVEVWMDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNA
QLVSGVYFQGTITIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFQGMNHDTLDRGCSCQMAVEKGGC
IMNASTGYPPFPMVFSSCSRKDLTSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEECDCEPEECMNRCNA
TTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGASPHCPANVYLHDGHSCQDQDGYCYNGICQT
HEQQCVTLWGPGAKPAPGICFERVNSAGDPYGNCGKVSFSSFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIET
NIPLQQGGRILCRGTHVYLGDDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRKNCHC
EAHWAPPFCDFKFGGGSTDSGPIRQAEARQEAESNRERGGQEPVGSQEHASTASLTLI
```

Signal peptide:
amino acids 1-28

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

90/615

FIGURE 89

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGG
CGCGGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTGACCTGAAAAAAATGTTCTGGATTTCTA
GAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTA
CTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCA
TACCATGCCTGTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGT
GATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTTCATGTTGGCCTTT
GGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATAACCTGGA
ATTGCTGTATTTTTCCAGAATGCCTTCATCTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGCCGCACTGAAGACTTA
TGGCAGTGAACACATCTGATTTCCACAGCACAACAGCCCTGCATGGGTTTGTGTTGTTTTTTTACTGCTCACTCC
CAACCTTTTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACTAAAATC
ACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCTATAAAAATGTTAGAGGAAAC
TTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATTTGTAAAAATAAAAAGAAATTACAAAAGAAAT
ATGGATTTGTCAATGTAAGTATTTGTCATATCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGT
GTTTATTCAAATGTGGTCTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTTAAATATTCC
GTGGTCAAATTTCTTCTCACTATAATTGGTATTTACTTTTACCAAAAATTTCTGTGAACATGTAATGTAACCTGGC
TTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCATGGTCCAGCCACCAGGCTCCCTG
TGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATTTCCAAGGGCAGGAAGATGTGACTCAGCCATGACACG
TGTTTCTGGTGGGATGCACAGTCACTCCACATCCACCACTG

1005556.0410.000000

91/615

FIGURE 90

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVIATIAFLMINAVSN
GQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEKDIVYPGIAVFFQNAFIFFGGLVFKFG
RTEDLWQ

2025-01-09 09:55:00

FIGURE 91

CGACGCCGGCGT**ATGT**GGCTTCCGCTGGTGTCTCCTGGCTGTGCTGCTGCTGGCCGTCCTCTGCAAAGTTTA
 CTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTCAAACGGCCCCCAGCGCCCCCTGGTAAC
 TGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGCTTTTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGT
 GGTAAATTGGCAGTGGCTTTGGGGGCCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCTCTGGTGTCT
 GGAACAACATACCAAGGCAGGGGGCTGTGTCTATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAATCCA
 TTACATTGGGCGTATGGAAGAGGGGAGCATTGGCCGTTTTATCTTGGACCAGATCACTGAAGGGCAGCTGGACTG
 GGCTCCCCTGTCTCTCTCTTTTACATCATGGTACTGGAAGGGCCCAATGGCCGAAAGGAGTACCCCATGTACAG
 TGGAGAGAAAGCCTACATTAGGGCCCTCAAGGAGAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAA
 GCTGGTTAAGGTGGTATCCAGTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTTCAGCT
 CCTCGACAGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGAGGTCTCT
 GCAGCAGTGGGGGCCCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCACTTACGGTGTCAACCCCAA
 CCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTACATGAAAGGAGGCTTTTATCCCCGAGGGGGTTC
 CAGTGAATTTGCCCTCCACACCATCCCTGTGATTTCAGCGGGCTGGGGGCGCTGTCTCACAAAGGCCACTGTGCA
 GAGTGTGTGCTGGACTCAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTA
 TTGCCCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGCCCGCTGCCT
 GCCAGGTGTGAAGCAGCAACTGGGGACGGTGCAGGGCCCGGCTTAGGCATGACCTCTGTTTTCTCTGCTGCGAGG
 CACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTATGTTTACTATGACACGGACATGGACAGGCGATGGA
 GCGCTACGTCTCCATGCCAGGGAAGAGGCTGCGGAACACATCCCTCTTCTTCTTCTGCTTTCCCATCAGCCAA
 AGATCCGACCTGGGAGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTT
 TGAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGGAGTACTATGAGACCTTCAAAAACCTCTTTGTGGAAGC
 CTCTATGTGAGTGGTCTGAAACTGTTCCCAAGCTGGAGGGGAAGGTGGAGAGTGTGACTGCAGGATCCCCACT
 CACCAACCACTTCTATCTGGCTGCTCCCCGAGGTGCCTGCTACGGGGCTGACCATGACCTGGGGCCGCTGCACCC
 TTGTGTGATGGCCTCCTTGAGGGGCCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACCTG
 TGGACTGGTTCGGGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTGTACTCAGACCT
 TAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAAT**TAGT**TCCATCAGGGAGGAGTCAGAGGAATTTG
 CCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTGCATTAGTTCCTTGCACGTATAAAGCACT
 CTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAGTTTAAATCACAATTCGAATCTGGGGCAATGGAATCACTGC
 TTCCAGCTGGGGCAGGTGAGATCTTTACGCCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATA
 GCTTGATGTCTCATGACGAGCGGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGAATATT
 CCATCTGTGGATAGAACCCCTGGCAGTGTGTGAGCTCAACCTGGTGGGTTCAGTTCTGTCTGAGGCTTCTGTCT
 CTCATTCAATTTAGTGTACGCTGCACAGTTCTACACTGTCAAGGGAAAAGGGAGACTAATGAGGCTTAACTCAA
 ACCTGGGGCTGGTTTGGTTGGCATTCATAGGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCAGTGGCT
 CTTCAGGGGACAGGAATGCCGTGTCTGGCCAGTTGTGGTGTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGT
 TAGTAGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTTATCATCGGGTG
 GCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGCCTTATCCACCAATACACAGG
 GAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCATGGACTGGTAAGATGAATACTTTGCTGGGCTGAA
 GCAGGCTGCAGGGCATTCCAGCCAAGGGCACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTC
 ACATCAGAAAAGGGAAAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTCAGTTAATTAGCACATGTGAG
 GGTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTGGTATCAGACAT
 ACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTGATTCCATTGCTTTAAAAA
 AAAAA

93/615

FIGURE 92

MWLPLVLLLVALLLAVLCKVYLGLFSGSSPNPFSSEVDKRPAPLVTDKEARKKVLKQAFSANQVPEKLDVVVIGS
GFGGLAAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLFDTGIHYIGRMEEGSI GRFILDQITEGQLDWAPLS
SPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQEEAIDKYIKLVKVSSGAPHAILLKFLPLPVVQLLDRC
GLLTRFSPFLQASTQSLAEVLQQLGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHMKGGFYPRGGSSEIA
FHTIPVIQRAGGAVLTkATVQSVLLDSAGKACGVS VKKGHEL VNIYCPIVVS NAGLFNTYEHLLPGNARCLPGVK
QQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTMDQAMERYVSMPREEAAEHIPLLFFAFPSAKDPTW
EDRFPGRSTMIMLIPTAYEWFEWQAE LKGKRGSDYETFKNSFVEASMSVVLKLF PQLEGKVESVTAGSPLTNQF
YLAAPRGACYGADHDLGRLHPCVMASLRAQSPINLYLTGQDIFT CGLVGALQGALLCSSAILKRNL YSDLKNLD
SRIRAQKKKN

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted November 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 93

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGGCAGCAGTGGCGGCGATGTTTGTGCGGCTCGGGATGG
 GTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCACAGCAAGTCGGGGCGGGTCAAACGTT
 CGAGTACTTGAACGGGAGCACTCGCTGTGGAAGCCCTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAA
 TCTGATGGGCAATGCCATGGTGATGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTT
 GTGGAACCGGGTGCCATGTTTCTGAGAGACTGGGAGTTGCAGGTGCACCTCAAATCCATGGACAAGGAAAGAA
 GAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGCCTGTGTTTGGAAACAT
 GGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAATGAGGAGAAGCAGCAAGAGCGGGTATTCCC
 CTACATCTCAGCCATGGTGAACAACGGCTCCCTCAGCTATGATCATGAGCGGGATGGGCGGCCCTACAGAGCTGGG
 AGGCTGCACAGCCATTGTCCGCAATCTTCATTACGACACCTTCCTGGTGATTGCTACGTCAAGAGGCATTTGAC
 GATAATGATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGGCCGGAGTCCGCCTGCCCGCGG
 CTACTACTTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCAATTTCTTGAAGTTGTTTGA
 ACTGACAGTGGAGAGAACCCCAAGAGGAAAAGCTCCATCGAGATGTGTTCTTGCCCTCAGTGGACAATATGAA
 GCTGCCCTGAGATGACAGTCCACTGCCGCCCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGT
 GTTTTCTGTATTTGCCATAGTCATTGGTATCATACCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTA
 CTGAGCCCTCCTGCTGCCACCACCTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACCTGGCCTGAGCA
 TGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTGTCACTGGAGTTTGAATGCAGG
 GACCCCGCATTTCCCATGGTTGTGCATGGGGACATCTAATCTGGTCTGGGAAGCCACCCACCCAGGGCAATGCT
 GCTGTGATGTGCCCTTCCCTGCAGTCCCTCCATGTGGGAGCAGAGGTGTGAAGAGAAATTTACGTGGTTGTGATGC
 CAAAATCACAGAACAGAATTTATAGCCAGGCTGCCGTGTTGTTGACTCAGAAGGCCCTTCTACTTCAGTTTTT
 GAATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTGTTGGGTTTTGCATTTGACC
 CAACCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACTTCTTCCCTGCCTTACCTTCCCTTTCA
 CTCCATTCAATTGTCTCTCTGTGTGCAACCTGAGCTGGGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGG
 CTGCAGAACACACCTGCGTTTCACTGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTG
 TTCCCTAGCATGGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTGAA
 GTTTGGCTAAAGGTTGGTGTAATAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATGGATTAGCTGTGCAAC
 TGACCAGCTCCAGGTTTGATCAAACCAAAGCAACATTTGTGTCATGTGGTCTGACCATGTGGAGATGTTTCTGGAC
 TTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGTTACGATTTTTTGAATCCCACTTTGAGTGTCTGAAAGTGTAAG
 GAAGCTTTCTTCTTACACCTTGGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTTCTTAATGGACAA
 GAGACAGTTGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCCTCATCATCTGTGCTTGGAGAGTTCACTG
 TCATTGAGCAGCACAGCCTGAGTGTGGCCTCTGTCAACCTTATTCCACTGCCTTATTTGACAAGGGGTTACAT
 GCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAGAGTCTCCTTGGAGGGCCTGGAACTCTGAGTC
 CTCCTATGAACCTCTGTAGCCTAAATGAAATCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCG
 GCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

[illegible]

FIGURE 94

Signal sequence:
amino acids 1-38

Transmembrane domain:
amino acids 310-329

[illegible]

97/615

FIGURE 96

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKA RSCGEVRQAYGAKGFS LADIPYQEIAGEHLRICPQ EYTCCTTEMEDKL
SQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMFVRTYGM LYMQNSEVFQDLFTELKRY Y
TGGNVNLEEMLNDFWARLLERMFQLINPQYHFSE DYLECVSKYTDQLKPF GDVPRKLKIQVTRAFIAARTF VQGL
TVGREVANRVSKVSPTPGCIRALMKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAE
RLEGPFNIESVMDPIDVKISEAIMNMQENSMQVSAKV FQCGQPKPAPALRSARSAPENFNTRFRPYNPEERPTT
AAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDES VTAGTSNEEECWNGH SKARYLPEIMNDGLTNQINNPEVD
VDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQD TSDESSGSGSGSGCMDDVCPT EFEFEFVTTEAPAVDPDRR
EVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

20050301 095200

FIGURE 97

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGATATTGACAAACTG
AAGCTTTCCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGACAGCTTTAGTGGCCGGCCGGCCGCTC
TCATCCCCCGTAAGGAGCAGAGTCCTTTGTACTGACCAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAA
TGGGAAGGTTTTATTGAAACTACAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTG
CAGAAATTTTATCCAACCTTTGTTTGGAAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTTCAT
AGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGCCATTCAAAGATGAATT
TCATTACGGTTGCGTTTTTAATCGGAGAGGACTGGTTGCCATGGCAAATGCTGGTTCTCATGATAATGGCAGCCA
GTTTTTCTTCCACTGGGTCGAGCAGATGAACCTTAACAATAAGCATAACCATCTTTGGAAAGGTTACAGGGGATAC
AGTATATAACATGTTGCGACTGTCAGAAGTAGACATTGATGATGACGAAAAGACCACATAATCCACACAAAAATAAA
AAGCTGTGAGGTTTTGTTTAAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGAGAAACC
AGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTTTCAATTTGGAGAGGAAGCTGA
GGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAGGGCAAAAGCAAAAGTAGTCATGACTTGCTTAA
GGATGATCCACATCTCAGTTCTGTTCCAGTTGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGG
AGAAGATGAAAGTGAGAGCATGATGAATATATGATGGTGATGAAAAAGAACCTGATGAGAGAAAGAAATTGCCAA
AAAATTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGAGAGAAGAAATCAGTCAGCCG
CAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACCTTAGCAGCAAAACAAAAAAGTAGAAAA
TGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAAGCCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAG
AGAAAAGCAAAAGTATGAAGCTTTGAGGAAGCAACAGTCAAAGAAGGGAACTTCCCGGGAAGATCAGACCCCTTGC
ACTGCTGAACCAAGTTTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGA
AGTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAAGCAGAAAAGTGAAAGATGCAAG
CATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGAATAAAAGAAGGAGGGAAGAAAGCAA
AAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAATGAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCT
ACAATGGCCTTGTAACAGCCATTGTTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTT
GTCTGGTTTTTGA AAAACAATTATCTTGTGTTTGC AAATGTGGAATGATGTAAGCAAAATGCTTTTGGTTACTGGTA
CATGTGTTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCTTCCACAAAAA
AAAAAAAAAAAAAAAAAAAA

99/615

FIGURE 98

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFIVQGGDPTGTGSGG
ESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELNKHTIFGKVTGDTVYNMLRLSEVDID
DDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEKPVEEVKKLKPCKTKNFSLLSFGEAEAEAEAEAEVNRVSQSM
KGKSKSSHDLKDDPHLSSVPVVESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSA
GEGEVEKKSVSRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEEAPPDGAVA EYRREKQKYEALRKQQS
KKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVKDASMQDSDTFEIYDPR
NPVKNRRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

20250610 09:25:00

CTTTTCTGAGGAACCCACAGCA**AATGA**ATGGCTTTTGCATCCTTGCTTCGAAGAAACCAATTTATCCTCCTGGTACTA
TTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTACCGCTGAAGTCTGTGCCACACACACA
ATTTACACGGACCCAAAGGAGATGATGGTGAAAAAGGAGATACGGAGGAAGAGGGGAAAAGCATGGCAAAGTGGG
CGCATGGGGCGGAAAGGAATTAAAGGAGAACTGGGTGAGATGCGAGATCAGGGCAATATTGGCAAGCTGGGCC
ATTGGGAAGAAGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAGCAGGTACTGTC
TGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCTCAAGACATCTATGAAG
TTTGTCAAGAAATGTGATAGCAGGGATTAGGGGAACTGAAGAGAAATTCCTACTACATCGTGCAGGAAGAGAAGAAC
TACAGGGAATCCCTAACCCACTGCAGGATTGCGGGGTGGAATGTCTAGCCATGCCCAAGGATGAAGCTGCCAACACA
CTCATCGCTGACTATGTTGCCAAGAGTGGCTTCTTTGCGGTGTTTCATTGGCGTGATGACCTTGAAAGGGGAGGA
CAGTACATGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGGAACCCAGCGACCCCTAT
GGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTGCCATCTTACCATGTACTTT
GTCTGTGATTGTTTATCAAGAAGAAAAAG**TAA**CTTCCCTCATCTACGTATTTGCTATTTTCTGTGACCGTCATTA
CAGTTATTGTTTATTCATCTCTTTTCTGATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAA
ACTGAGGTATGGAGCCTCCATCATCAAAAAA

101/615

FIGURE 100

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEGKHGKVGRMGPKGI
KGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKKGAGTVCDGGRYRKFGQLDISIARLKTSMKFVKNVIA
GIRETEEFYYIVQEEKNYRESLTHCRIRGGMLAMPKDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDN
TPLQNYSNWNEGEPSDPYGHEDCVEMLSGRWNDTECHLTMYFVCEFIKKKK
```

Signal peptide:
amino acids 1-25

101/615

FIGURE 101

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGACCCAGAGCTTCTCC
 AGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCCTTAACCTTCCCTCCGCGGGGCCAGCCACCTTCGGGAGTCCGGGTT
 GCCCACCTGCAAACCTCTCCGCCCTTCTGCACCTGCCACCCCTGAGCCAGCGCGGGCCCCCGAGCGAGT**CATG**GCCA
 ACGCGGGGCTGCAGCTGTTGGGCTTCATTCTCGCCTTCCCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGC
 CCCAGTGGAGGATTTACTCCTATGCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGATGT
 CCTGCGTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGC
 AAGCAACCCGTGCCTTGATGGTGGTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
 AGTGTATGAAGTGCTTGGAGACGATGAGGTGCAGAAGATGAGGATGGCTGTTCATTGGGGGTGCGATATTTCTTC
 TTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCCTATGA
 CCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCTTCTGCTG
 GAGGTGCCCTACTTTGCTGTTCCCTGTCCCCGAAAAACAACCTCTTACCCAACACCAAGGCCCTATCCAAAACCTG
 CACCTTCCAGCGGGAAGACTACGT**TGAC**ACAGAGGCAAAAGGAGAAAAATCATGTTGAAACAAACCGAAAAATGG
 ACATTGAGATACTATCATTAAACATTAGGACCTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAAC
 AAACAAACAAACAAAAAACCCATGTGTTAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTT
 CTCAATATAGGAGGGAAGATTTTCCATTTGTATTACTGCTTCCCATTTAGTAATCATACTCAAATGGGGGAAGG
 GGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAAATAGACAGTAAAATACTATTCT
 CATTATGTTGATACTAGCATACTTAAATATCTCTAAATAGGTAAATGTATTTAATTCCATATTGATGAAGATG
 TTTATTTGGTATATTTTCTTTTTCGTCCTTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGC
 TTTGGGTGCCCTTTGCCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTT
 TCATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCCATCGTTATTAAGCCCTTATTTGTTTTGTGTT
 TCATTGGTCTCTATCTCCTGAATCTAACACATTTTCATAGCCTACATTTTAGTTTCTAAAGCCAAGAAGATTTAT
 TACAAATCAGAACTTTGGAGGCAAACTTTTCTGCATGACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAAT
 CCCTGTACTCTGACCATAGCACTCTTGTTTTGCTTTGAAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCC
 CCAGGTGTTGTAACACAACCTTATTTGATTGAATTTTAAAGCTACTTATTCATAGTTTTATATCCCCCTAAACTAC
 CTTTTTGTTCCTTCCCTTAAATTTGATTGTTTTTCCCAAGTGTAATTATCATGCGTTTTATATCTTCTTAATAAG
 GTGTGGTCTGTTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATAATCTGGTGACAAATATTTCTCTGTAGC
 TGTAAGCAAGTCACTTAATCTTTCTACCTCTTTTTTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAG
 AAGAGGTAGTGTAATATTAATTAGTTTATATTACTCTTATTCTTTGAACATGAACATATGCCTATGTAGTGTCTT
 TATTTGCTCAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATTCACTG
 CCTTCTCTCTCTACAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCATGTGGTTTCACTGCCTTCTCT
 CTCTTACCAGTCTATTTCCACTGAACAAAACCTACGCACATACCTTCATGTGGCTCAGTGCCTTCTCTCTCTA
 CCAGTCTATTTCCATTCTTTTCACTGTGTCTGACATGTTTTGTGCTCTGTTCCATTTTAAACAACTGCTCTTACTTT
 TCCAGTCTGTACAGAATGCTATTTCACTTGAGCAAGATGATGTAATGGAAAGGTGTTGGCACTGGTGTCTGGAG
 ACCTGGATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGGAGCAAGGCATTTGGCTGCTGTAAGCTTATTG
 CTTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCACCTCACAGTGATGTTGTGGGGATCCAGTGAGATA
 GAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTATACTAAGGGAAAGAAATTGAGGAATTAAGTGCATAC
 GTTTTGGTGTGCTTTTCAAATGTTTGAATAAAAAAATGTTAAG

CCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCCAGCCTGCAGGGCTGATAAGCGAGC
CATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGCGCGCAGTAGAGCAGCAGCACAGGCGC
GGGTCCCGGGAGCGCGGCTCTGCTCGCGCCGAGATGTGGAATCTCCTTCACGAAAACCGACTCGGCTGTGGCCACC
GCGCGCCGCGCGCTGGCTGTGCGCTGGGGCGCTGGTGTCTGCGGGTGGCTTCTTCTCTCGGCTTCTCTCTTC
GGGTGGTTTTATAAAATCCTCCAATGAAGCTACATAACTTACCCAAGCATATAATGAAAGCATTTTTGGATGAA
TTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGCAGGAACAGAACAAAAC
TTTCAGCTTGCAAAGCAAATTCATCCAGTGGAAGAATTTGGCCTGGATTCTGTTGAGCTAGCTCATTATGAT
GTCCCTGTTGTCTACCCAAATAAGACTCATCCCAACTACATCTCAATAATTAATGAAGATGGAATGAGATTTT
AACACATCATTATTTGAACCACCTCTCCAGTATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGTCTTC
TCTCCTCAAGGAATGCCAGAGGCGCATCTAGTGATGTTAACTATGCACGAATCGAAGACTTCTTTAAATTTGGAA
CGGGACATGAAAATCAATTGTCTTGGGAAAATTGTAATTGCCAGATATGGGAAAGTTTTTCAGAGGAAATAAGGTT
AAAAATGCCCAGCTGGCAGGGGGCCAAAGGAGTCATTCTCTACTCCGACCTGCTGACTACTTTTGCTCCTGGGGTG
AAGTCTATCCAGACGGTTGGAATCTTCTGGAGGTGGTGTCCAGCGTGGAATATCTAAATCTGAATGGTGC
GGAGACCCTCTCACACCAGGTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTT
CCAAGTATTCCTGTTTCATCCAATTGGATATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGCTCAGCACCA
CCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTCCCTACAATGTGGACCTGGCTTTACTTGGAACTTTTTCTACA
CAAAAAGTCAAGATGCATCCACTCTACCAATGAAGTCAGAGAATTTACAATGTGATAGGTACTCTCAGAGGA
GCAGTGGAACCAGACAGATATGTCAATCTGGGAGGTCAACGGGACTCATGGGTGTTTGGTGGTATTGACCCCTCAG
AGTGGAGCAGCTGTTGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGA
ACAAATTTTGTGTTGCAAGCTGGGATGCAGAAGAAATTTGGTCTTCTTTGGTTCTACTGAGTGGGCAGAGGAGAATTCA
AGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTATAGAAGGAACTACACTCTGAGAGTT
GATTGTACACCTGTATGTACAGCTTGGTACACAACCTAACAAAAAGAGCTGAAAGGCCGTGATGAAGGCTTTGAA
GGCAAATCTCCTTATGAAAGTTGGACTAAAAAAAGTCTTCCCAAGAGTTCAGGTGGCTAGCCGAGATAAGCAAA
TTGGGATCTGGAATGATTTTGAAGGTGTTCTTCCAACGACTTGAATTTGCTTCAGGCAGAGCACGGTATACTAAA
AATTGGGAAACAAACAAATTCAGCGCTATCCACTGTATCACAGTGTCTATGAAACATATGAGTTGGTGGAAAAG
TTTTATGATCCAATGTTTTAAATATCACCTCACTGTGGCCAGGTTCCAGAGGGGATGGTGTTTGAGCTAGCCAAT
TCCATAGTGTCTCCCTTTTGATTGTGAGATTATGCTGTAGTTTTTAAGAAAGTATGCTGACAAAAATCTACAGTATT
TCTATGAAACACTCCACAGGAAATGAAGACATACAGTGATCATTTGATTCATTTTTCTGCTAGCTAAAGAAATTT
ACAGAAATTTGCTTCCAAGTTCAGTGAGAGACTCAGGACTTTGACAAAAGCAACCCCAATGATATTAAGAAATGATG
AATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGACAGGCCTTTTTTATAGGCAT
GTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTCATTCCAGGAATTTATGATGCTCTGTTTGAT
ATTGAAAGCAAAGTGGACCCCTCCAAGGCCTGGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTG
CAGGCAGCTGCAGAGACTTTGAGTGAAGTAGCCCTAAGAGGATTTTTTATAGAGAATCCGTATTGAATTTGTGTGGTA
TGTCACTCAGAAAGAATCGTAATGGGTATATTGATAAAATTTTAAAATTGGTATATTTGAAATAAAGTTGAATATT
ATATATAA

105/615

FIGURE 104

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLLHETDSAVATARPRWLCAGALVLAGGFLLGFLFGWFIKSSNEATNITPKHNMKAFLDELKAENIKKFLH
NFTQIPHLAGEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTHPNYISIIINEDGNEIFNTSLFEPPPPG
YENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFFKLERDMKINCSGKIVIARYGKVFRGNKVNAQLAGAKGV
ILYSDPADYFAPGVKSYPDGWNLPGGGVQQRGNILNLNGAGDPLTPGYPANAYARRGIAEAVGLPSIPVHPIGYY
DAQKLEKMGGSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYINVIGTLRGAVEPDYVILG
GHRDSWVFGGIDPQSGAAVHEIVRSFGLTKKEGWRPRRTILFASWDAEEFGLLGSTEWAEENSRLQERGVAYI
NADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYESWTKKSPSPEFSGMPRISKLGSGNDFEVFF
QRLGIASGRARYTKNWETNKFSGYPLYHSVYETYELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDICRDY
AVVLRKYADKIYSISMKHPQEMKTYSVSFDLSFAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAF
IDPLGLPDRPFYRHHVIYAPSSHKNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAETLSEVA

Signal sequence:
amino acids 1-40

N-glycosylation sites.
amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340, 459-463,
476-480, 638-642

Tyrosine kinase phosphorylation sites.
amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.
amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341, 360-366,
427-433, 529-535, 707-713

105/615 99525001

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTTTACTGCTGCGTTTTATGTTGGGAATTCCTCTCCTATGGCCTTGTCCTTGAGACAAAGAACTCTCAAACAAAGAAAGTCAAGCAGCCAGTGCGATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTTTTTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGCCAGCTAAGATCTGATTTAGACAATGGAACAATTTCTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTTATCATTTGATGAAGAACAGGTGACATATGCCATACAGAAGCTTGATAGGAGGAGCATCCCTCTACATCTTAAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGAGTTTGTCTATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATGAGGCCATTGTACCAGAGATGTCTCCAGAGGAACATTAGTTATCCAGGTGACAGCAAGTGATGCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTAGCTTACTTCAAGGCCAGCCATATTTTTCTGTTGAACCAACAAAGGAGTCAATAAGAAATCTTCTTAAATGGATAGAGAACTGCAAGATGAGATGATTTGGGTAAATCATTTCAAGCCAAAGGACATGATTTGGTCAAGCCAGGCGTTGTCTGGACAACAGTGTATTAATTAACCTTTTCAGATGTTAATGACAATAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGATGATTGCAAAACATTTGACATTATTACTAATCATGAAACTCAAGAAAGGAATAGTTATATTTAAAAAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATCATGTTCTCTGAGCAGCTCATGAAGTACCACACTGAGGCTTTCCACCACCTTTTCATTAAGATCCAGGTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTAATTTGAAGTTTTGAAGAAACCCCAAGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAATCTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCATATCAATGATAATGGTACAATCACTACAAGTAACTCACTGGATCGTGAAATCAGTGCCTTGGTACAACCTAAGTATTACAGCCACAGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCAATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTTGTGAAAAATGCAAGTCTCTGGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGACCACTTTTACTTTAATCTATCTGTAGAAGACACTAACAAATTCAGGTTTTTACAATCATAGATAATCAAGATAACACAGCTGTCAATTTTGACTAATAGAAGTGGTTTTTAACTTCAAGAAAGAACCTGTCTTCTACATCTCCATCTTAATTTGCCGACAATGGAATCCCGTCACTTACAAGTACAACACCCCTTACCATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTGTGCTTTCCATGGGATTCAAGACAGAAGTTATCAATTGCTATTCTCATTTGCATTATGATCATATTTGGGTTTATTTTTTTTGACTTTTGGGTTTAAAAACAACGGGAGAAAAAGATTTCTATTTCTTGAGAAAAGTGAAGATTTTCAGAGAGAAATATATTCAAATATGATGATGAAGGGGTGGAGAAGAAGATACAGAGGCTTTTGATATAGCAGAGCTGAGGAGTAGTACATAATGCGGGAACGCAAGACTCGGAAAAACCAAGCGCTGAGATCAGGAGCCATATACAGGCAGCTTTTGCAGGTTGGCCCCGACAGTGCCATATTCAGGAAATTCATTCTGAAAAAGCTCGAAGAAGCTAATACTGATCCGTGTGCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAAACAGGTCATTAGCTGGATCCCTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTTGGGACCTCGCTTTAAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCGAGTCAAATAATTAGGGCTTTTTTACCATCAAAATTTTTTAAAGTGCTAATGTGTATTTCGAACCAATGCTAGTCTTAAAGAGTTTTGTGCCCTGGCTCATATGGCGGGAAAGCCCTAGTCTATGAGGATTTTCTGATTTCCCTGGAGTAAATATCCATGGTTATTTTAAAGCTACCTACATGCTGTCTGTTGAACAGAGATGTGGGGAGAAATGTAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAAATAATGTAGGAAGATATTTAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCCTTATGCGATTATATCATTTACTTTAGGAAAGAGTAAAAATACCAACGAGAAAAATTTAAAGGAGCAAAAAATTTGCAAGTCAAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAAAATGAAAAATGTATAGTCAGAGAAATTTTCATGAATTTATCCATGAAGTATTGTTTCTTTATTTAA

107/615

FIGURE 106

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGVWVNQFFVPEEMNTTSHHIGQLRSDLDNGNN
SFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIATGRAVEPESEFVIKVSINDNEPKFLD
EPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARLLYSLLQGQPYFSVEPTTGIVIRISSKMDRELQDEYWVIIQ
AKDMIGQPGALSGTTSVLIKLSDVNDNKPIFKESLYRLTVSESAPTGTSGITIMAYDNDIGENAEMDYSIEEDDS
QTFDIIITNHETQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLLPYY
VFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISAWYNLSITATEKYNIEQ
ISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDRDESIEHHFYFNLSVEDTNNSSFTIIDNQ
DNTAVILTNRGTGPNLQEEPVFYISILIADNGIPSLTSTNTLTIHVCDGDSGSTQTCQYQELVLSMGFKTEVIA
ILICIMIIFGFIFLTLGLKQRRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTS
AEIRSLYRQSLQVGPDSAI FRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESASVSDQDESYDYL
NELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ATCTGGTTGAACACTACTTAAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCATGATTTGACTCAGAGATT
CTCTTTTGTCCACAGACAGTCAATCTCAGGGGCAGAAAGAAAAGAGCTCCCAAATGCTATATCTATTTCAGGGGCTC
TCAAGAACAATGGAATATCATCCTGATTTAGAAAAATTTGGATGAAGATGGATATACTCAATTACACTTCGACTCT
CAAAGCAATACCAGGATAGCTGTTGTTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCTTGGCGCCTCATTTGCT
GTAATTTGGGAATCCTATGCTTGGTAATACTGGTGATAGCTGTGCTGCTGGTACCCTGGGGTCTTTTCCAGC
CCTTGCTCTCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTAGCATGTCACTAAATTCCTGGGATGG
AGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAGACAGCTCAAATGAATTGGGATTTATAGTA
AAACAAGTGTCTTCCCAACCTGATAATTCAATTTGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTC
TGGGAGGATGGATCAACATCTCTCTTCTAATCTATTTCAGATCAGAACCACAGCTACCCAAAGAAAACCCATCTCCA
AATTGTGATGGATTACAGTGTCACTGATCTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGTGAGAAG
AAGTTTTCAATGTAAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTAAGGAGGACAGAAAACAGAA
CAGAAAAGAGTAACAGCTGAGGTCAAGATAAAATGCAGAAAATGTTTAGAGAGCTTGGCCAACTGTAATCTTAACC
AAGAAATTGAAGGGAGAGGCTGTGATTTCTGTATTTGTGACCTACAGGTAGGCTAGTATTTATTTTTCTAGTTAG
TAGATCCCTAGACATGGAATCAGGGCAGCAAGCTTGAGTTTTATTTTTTATTTATTTATTTTTTGTAGATAGG
GTCCTACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCGAGTATCTCTCGCCCTCAGCCCC
TCAAGTAGCTGGGACTACAGGTGCATGCCACCATTGCCAGGCTAATTTTTGGTGTTTTTGTAGAGACTGGGTTTT
GCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAGTGATCTGCCCGCTTGGCCTCCCAAAGTGTCTGGGA
TTACAGATGTGAGCCACCACACCTGGCCCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTTACCTTGGG
TAAGCCATAAGCGAATCTTAATTTCTGGCTTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGACACGG
TGTGTTGCCACGATTTGACCCCTCAACTCTTAGCAGTATATCAGTTATGAATGAGGGTGAAATATATTTCTGAAT
AGCTAAATGAAGAAATGGGAAAAAATCTCACCAAGTGCAGGCAATTTTATTTTTCATCAGTATGATACATAA
TTATGATTATCATCTTAGTAAAAAAGCAGGAACCTCTACTTTTTCTTTATCAATTAATAGCTCAGAGAGTACATC
TGCCATATCTCTAATAGAATCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTTTCGCTCTTGTGTTGCCAGGCTG
GAGTGCACACGGCAGATCTCGGCTACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCC
CAAGTAGCTGGGATTACAGTCAAGGCACCAACACACCCGGCTAATTTTGATTTTTTTTAGTAGAGACAGGTTTCT
CCATGTCGGTCAGGTTAGTCCCAGACTCCTGACCTCAAGTGATCTGCCCTGCCCTCGGCCCTCCCAAGTGCTGGGATT
ACAGGCGTGAGCCACTGCACCCAGCCTAGAATCTTGTAATAATGTAATTGTAGGGAAACTGCTCTCATAGGAAA
GTTTTCTGCTTTTTTAAATACAAAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACC
TCATTGGAACAAGTATTAACATTTTGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTTTACCATT
TTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAAACTATATTTCTCATGTGCTGATTGTCTTATTT
TTTTTCATATCTTCCCACCTGGTGCIATTTTTTTTCCAAATGGATATTTCTGTATTACTAGGGGACCATTTACAGTC
CTCTAATGTTGATTAAATATGTGAAAAGAAATGTACCAATTTTACTAAATATGCAAGTTTAAATGGATGATTTT
ATGTTATGTGATTTTCAATTTCAATAAAAAATACTCTTATCAAAAAAATAAAAAAATAAAAAAATAAAAAA

FIGURE 108

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDEGYTQLHFDSQSNTKRVAVSEKGS CAASP PWRLIAVILGILCLVILVIAVVLGTMGVLLSSPCP
PNWIIIEYKSCYLFMSLNSWDGSKRQCWQLGSLNLLKIDSSNELGFIVKQVSSQPDNSFWIGLSRPQTEVPWLWED
GSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQLCSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 109

CTGCAAGTTGTTAACGCCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCACA
 ATATCTTAACCTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAGAC
 CTACCTACCCGTACGCATACATACATATGTGTATATATATGTAACTAGACAAAGATCGCAGATCATAAAGCAAG
 CTCTGCTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAGATGTTATTTGTCAAGATCCCTGTCGATTTCATGCCCT
 TTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCTTTGGTTTTGGGGACATTATGATTTGTGTAAAGACTCAGAT
 TTACACGGAAGAAGGAAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATATCTGAA
 AGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCCTCTGAGACGTTCTGTGCAATGGGCAATCCCTACAT
 GTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGAATTTTGAAGGAAG
 ACATCCCTCCACATTTTGGCAGTCTGCCACTTGGGAAGGAGTATCCCAAGCCTCTCCAGGTAAACATCACTCTGTC
 TTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACCAAATGATCCT
 GGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTAGATGCTTTTCA
 CATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTCACAGAAGAGTACTCAAC
 AGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTTCGCGCTTTTGTCTGGACCTCGCCT
 ACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAAACCAAGAACTCAGAGATTTCTTTACAGTCACAGACCT
 GAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGCACGCTACTTTTACGC
 GATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGAATCTCCATGCCACTGTATGTGTGTATGACAACAGCAA
 ATTGACATGCGAATGTGAGCACAACTACAGGTCCAGACTGTGGGAAATGCAAGAAGAATTATCAGGGCCGACC
 TTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAAATACCTGTATCCCCAGTATTTCCAGTATTGG
 TACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGAGGGACGTGCCACAACAACGTGCGCTGCCTGTG
 CCCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGCAGCTGCGGCTCCGACTCTGG
 CCAGGGCGCGCCCCCGCACGGCACCCCGAGCGCTGCTGCTGCTGACCACGCTGCTGGGAACCGCCAGCCCCCTGGT
 GTTCTAGGTGTACCTCCAGCCACACCGGACGGGCCCTGTGCCGTGGGGAAGCAGACACAACCCAAACATTTGCTA
 CTAACATAGGAAACACACATACAGACACCCCCACTCAGACAGTGTACAACTAAGAAGGCCTAACTGAACTAA
 GCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTCTGACTCCAGAGGAGTTGGCAGCTG
 TTGATATTTATCACTGCAAAATCACATTGCCAGCTGCAGAGCATATTGTGGATTGGAAAGGCTGCGACAGCCCCCA
 AACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTGGCTACTCTAGCGTGGTGCGCCCTAGTAC
 GACTCCGCCCAGTGTGTGGACCAACCAATAGCATTCTTTGCTGTGAGGTGCATTGTGGGCATAAGGAAATCTGT
 TACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAACCTGTGCTTTAGTGAACGTTGCTCTGTAA
 CCCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACATGTGTAACAGCCCCCTCTAAAAGCGCAAG
 CCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGAGCACACCCCACTATACAAGAGTGGCTATA
 GGAAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTTATTTTCTTGAACACTGTAATATGTAGATT
 TTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAATGTATCTAATTCGAATCAGCAAAGACTGACATT
 TTATTTTGTCTCTTTCTGTTCTGTTTTGTTTTCACTGTGCAGAGATTTCTCTGTAAGGGCAACGAACGTGCTGGCA
 TCAAAGAATATCAGTTTACATATATAACAAGTGTAATAAGATTCCACCAAAGGACATTCTAAATGTTTTCTTGT
 GCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCATAAACGATTTTCAGGAATTTGTATTGCAATTTCTTA
 AGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACTTTACTGATTTCTGTGTGGACTGAGTACATTAGCTG
 ACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACTAGCTTGGACAACCTCTGCAAAATATGAGACTATTTCC
 ACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAAAAAAAAAA

111/615

FIGURE 110

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLKVKLDPPDITCGDP
PETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWKEYPKPLQVNITLSWSKTIELTDNIVI
TFESGRPDQMLEKSLDYGRTWQPYQYYATDCLDAFHMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEI
KDRFALFAGPRLRNMASLYGQLDTTKKLRDFFTVDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCN
LHATVCVYDNSKLTCECEHNTTGPD CGKCKKNYQGRPWS PGSYLPIPKGTANTCIPSISSIGTNVCDNELLHCQN
GGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTASPLVF

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

112/615

FIGURE 111

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCTCGCCCTGCACCGCGTAGACCGAC
CCCCCCTCCAGCGCGCCCAACCGGTAGAGGACCCCGCCCGTGCCCGACCGGTCCCCGCCCTTTTGTAAAAC
TAAAGCGGGCGCAGCATTAACGCTTCCCGCCCCGGTGACCTCTCAGGGGTCTCCCGCCAAAGGTGCTCCGCCG
TAAGGAACATGGCGAAGGTGGAGCAGGTCTCTGAGCCTCGAGCCGACGACGAGCTCAAATTCGAGGTCCCTTCA
CCGATGTTGTACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTAAAGGTGAAGACTACAG
CACCACGTAGGTACTGTGTGAGGCCAACAGCGGAATCATCGATGCAGGGGCCCTCAATTAATGTATCTGTGATGT
TACAGCCTTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTTTATGGTTCAGTCTATGTTTGTCTCCAACTG
ACACTTCAGATATGGAAGCAGTATGGAAGGAGGCAAAACCGGAAGACCTTATGGATTCAAAACTTAGATGTGTGT
TTGAATTGCCAGCAGAGAATGATAAACCATGATGTAGAAATAAAATAAAATTATATCCACAACCTGCATCAAAGA
CAGAAACACCAATAGTGTCTAAGTCTCTGAGTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGGAAGAAT
GTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAAGCAGTTCAAGGAAGAAGATGGACTGCGGA
TGAGGAAGACAGTGCAGAGCAACAGCCCCATTTACGATTAGCCCCAACTGGGAAGGAAGAAGGCCCTTAGCACCC
GGCTCTTGGCTCTGGTGGTTTTGTTCTTTATCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAGAGGTAGCATG
CACAGGATGGTAAATTGGATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGA
AATTAATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACACAGATACACA
CACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTGATTGAGGGGGAAAAAGAATGAT
CTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGG
CCTTGGTACATGATGCTGGATTACCTCTCTTAAATGACACCCCTTCCCTCGCCTGTTGGTGTGGCCCTTGGGGAG
CTGGAGCCCAGCATGCTGGGGAGTGCGGTGAGTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCAGGCT
GCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGTCTCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAA
TTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGA
AATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTG
TTATTTCAGAGATGTTTAAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTACAAAGAGT
ACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCTGT
CTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCAGGAAGTGTTTTTTCT
GGGTCAGTAAATAACAACCTGTCTAGGGAGGGAAATTCTCAGTAGTGACAGTCAACTCTAGGTTACCTTTTTTAA
TGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATACCACCTCTCAACCATTACTCACACTTCCAGCGCCCAAGGTCC
AAGTCTGAGCCTGACCTCCCCCTTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAA
GCGAGGGCACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTACTAGTT
GAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAA
AGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

1005556 01503

FIGURE 112

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFPTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIIDAGASINVSVMQLQP
FDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRCVFELPAENDKPHDVEINKIISTTASKTET
PIVSKLSSGLDDTEVKKVMEECKRLQGEVQRLREENKQFKEEDGLMRKTVQSNSPISALAPTGKEEGLSTRLL
ALVVLFFIVGIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 113

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGGCGGGTAGCTGGCA
GGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCTCATGACTTCTCTTGTGGACC**ATG**TCCGTGATCTT
TTTTGCCTGCGTGGTACGGGTAAAGGGATGGACTGCCCCCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTT
TTTGGAATGGAGGAGACGGCTCAAGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGG
TTGTGACTTTAGTATACATTTTTCTTCTTTCTGGGACGTTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCCAGC
AGCCATGGCCCTTCTGCTTCCCTGGAGACCCCTGTGGTGGGAATTACAGCTTCTATGACACTACCTGCATTGGCCCT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATTTTAACTATGTAAG
TTCTCTCAGATGGAGTGCAGCTTGAAAAAATTACAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCT
GGAGGACACAGATGTGGCAAATGGGGTGATGAATGGTCACACACCCGATGCACCTTGGAGCCTGCTCCTAATTTCCG
AATGGAACCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCTCATTCTG
AGGAGTTCACCTTGACAGAACATTTCTTACAGGATCCAAGGAGCTGGTTCTGCTGGTTGGACCAAACCTCG**TGAGC**
CAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCATCCGGGAGCAGTGATGTCAAACCTTCTGCTGCTG
GGGAAATCTCATCAGCAGGGAGCCTGTGGAAAAGGGCATGTGAGTGAATCTGGGAATGGCTGGATTCGGAAACA
TCTGCCCATGTGTATTGATGGCAGAGCTGTTGCCCAAGCGCCTTTTATTTAGGGTAAAAATTAACAAATCCATT
CTATTCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTTGCTTCAGAA
GTGTTATTTTCATGAATCATTATATGATTTGATCCCCCAGGATTCTATTTTGTTTAATGGGCTTTTCTACTAAAA
GCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCATTTACTTTACATATTCGTTTTCAATACTTGCTGTTC
ATGTTACACAAGCTTCTTACGGTTTTCTTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAAACAAAC
TCAGTAGTACAACCTAACTTGTATAAAAGTGTGTAATAAATGTATAGCCATTTATATCCTATGTATAAATTAAT
GAGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAACCAAAAAAAAAAAAAAAAAAAAAAG

115/615

FIGURE 114

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHFSSFGDVACMAICS
CQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQVKWHFNYVSSSQMECSLEKIQEELKLQPP
AVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTALGILSLIINIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

2025.01.15 10:45:00

CTCAGCGCGCTTCTCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGACCCGACCTTAAAG
AGTGGGGAGCAAAGGGAGGACAGAGCCTTTAAAAACGAGGCGGGTGGTGCCTGCCCTTTAAGGGCGGGGCGTCC
GGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTCTGTGCGAGGCTGCGAGGAAAGGCCCTTAGGCTGGGTC
TGGGTGCTTGGCGCGCGGCTTCTCTCCCGCTCGTCTCTCCCGGCCAGAGGACCTCGGCTCATGCT
GAGCAGAGTA**TGGA**AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGGATCCGCGAG
TGTTATTATATCAACACTTCTGTTTGAACACTGTACATCCTCTGCCACATCTTCTGACCCGCTTCAAGAAGCCT
GCTGAGTTACACACAGTGGAATGATGAAGATGCCACCGTCAACAAGATTGCGCTCGAGCTGTGCACCTTTACCCTG
GCAATTGCCCTGGGTGCTGTCTCTCTCTCTCTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCCTCGG
AATCTACTACATCCGATGGCTCAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTG
TCCCTCATCTTCTCTCATGCCCTTGGCATATTTCTTCACTGAGTCTGAGGGCTTTCGTGGCTCGAGAAAGGGTGTC
CTGGGCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCTAGGTATGGTGTTGGTGCCA
TCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCTATGACTTTTTGGGAGTACTATCTCCCTACCTC
TACTCATGCACTCTCTTCTTGGGGTCTGCTGCTCCTGGTGCTACTCCACTGGGTCTCGCCCGCATGTTCTCC
GTCACCTGGGAAGCTGTAGTCAAGCGCTCGGTCTGGAAGCCTGGAGGAGGACAGCTGTAGCTCAGCCTTTGAG
GAGGCGCCCTGACCCGAGGATCTGTAATCTACTTCTGCTGGCTGCCCTTTAGACTGAGCTGCTACACAGA
CAGGTCTTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGGCTTCAGCCTGGCAACGGAACCTG
GGTACCCCTTGGCTATGCTGTGCTTGTGCTGGTGCTGACGGGCCTGTCTGTGCTCATTGTGGCCATCCACATCCTG
GAGCTGCTCATCGATGAGGCTGCCATGCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCTTCTCCAAG
CTGGGCTCCTTTGGTGGCGGTCACTCAGGTTGTACTCATCTTTTACCTAATGGTGCTCTCAGTTGTGGGCTCTAT
AGCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACATGCCATGCCATGACGCAGATAAATGGGAAGCTGTGTC
TGCTCTCTGGTCTTAAGCTCAGCACTTCTGTCTTCTCTCGAACCTGGGGCTCACTCGCTTTGACCTGCTGGGT
GACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACATTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACC
ACACTCTGCTGGTGGAAGACCTTCACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTG
CCGCTGCCCGTCTCCGGTTTTCCCCAGGCATCTAGGAAGACCCAGCACCA**TGA**CCTCCAGCTGGGGGTGGGAAG
GAAAAAAGTGGACATGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGGACCTCAGGACCTGGA
ATCTGAGAGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCAGTAATTCGATAATCTGAGCCAGAGTTTGGGACCA
GGACCTCCTGCTTTTCCATACTTAACTGTGGCCTCAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCGTGAT
CCCAAATCTGTTTACACATCAATCTGCCTCACTGCTGTTCTGGGCCATCCCATAGCCATGTTTACATGATTGA
TGTGCAATAGGGTGGGGTAGGGGCAGGGCAAGGAAGGACTGGGCGAGGGCAGGCTCGGGAGATAGATTGTCTCCCTTGC
CTCTGGCCCCAGCAGAGCCTAAGCACTGTGCTATCTTGGAGGGGCTTTGGACCACCTGAAAGACCAAGGGGATAGG
GAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCAGGGAAAAAAA

117/615

FIGURE 116

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNKIALELCTFTLAIA
LGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSLIFLMPFAYFFTESEGFAGSRKGVLR
VYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESLYDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTG
KLLVKPRILLEDLEEQLYCSAFEEAALTRRICNPTSCWLPLDMELLHRQVLALQTORVLLLEKRRKASAWQRNLGYP
LAMLCLLVLTGLSVLIVAIHILELLIDEAAMPGRMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGIFYSSP
LFRSLRPRWHDAMTQIIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRFNWLGNFYIVFLYNAAFAGLTTLCL
LVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

121/615

FIGURE 120

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737
><subunit 1 of 1, 1141 aa, 1 stop
><MW: 124671, pI: 5.82, NX(S/T): 5
MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQLQPRPQSWLLVGAP
QALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWLGVSVRSQGPGGKIIVTCAHRYEARQRV
DQILETRDMIGRCFVLSQDLAIRDELDDGGEWKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGT
ARVELCAQGSADLAHLDDGPYEAGGEKEQDPRLIIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVI
LRKDSASRLVPEVMLSGERLTSGFGYSLAVADLNSDGWPDLIIVGAPYFFERQEELGGAVYVYLNQGGHWAGISPL
RLCGSPDSMFGISLAVLGLDLNQDGFDPDIAVGAPFDGDGKVFYIYHGSSLGVVAKPSQVLEGEAVGIKSFYSLSGS
LDMDGNQYPDLLVGLSLADTAVLFRARPILHVSHEVSIAPRSIDLEQPNCAAGHSVCVDLRVCFYSYIAVPSSYSPT
VALDYVLDADTDRLRGQVPRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYS
LQTPRLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVD
GTTALFALSGQPVIGLELMVTNLPSDPAQPPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSNENASHV
ECELGNPMKRGAQVTFYLLSTSGISIEETTELEVELLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFS
GVVRGERAMQSERDVGSKVKEYEVTVSNQGSRLTLGSAFLNIMWPHEIANGKWLLYPMQVELEGGQGPGQKGLCS
PRPNILHLDVDSRDRRRRELEPPEQQEPGERQEPSMSWVPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAA
VLHVWGRLWNSTFLEEYSVKSLVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMVAVAEVGPWWVILLAVL
AGLLVLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKKEKTGTILRNNWGSPPREGPDAHPILAAD
GHPELGPDPGHPGPGTA
```

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100																																																																																																																				
Population	122,741	123,155	123,569	123,983	124,397	124,811	125,225	125,639	126,053	126,467	126,881	127,295	127,709	128,123	128,537	128,951	129,365	129,779	130,193	130,607	131,021	131,435	131,849	132,263	132,677	133,091	133,505	133,919	134,333	134,747	135,161	135,575	135,989	136,403	136,817	137,231	137,645	138,059	138,473	138,887	139,301	139,715	140,129	140,543	140,957	141,371	141,785	142,199	142,613	143,027	143,441	143,855	144,269	144,683	145,097	145,511	145,925	146,339	146,753	147,167	147,581	147,995	148,409	148,823	149,237	149,651	150,065	150,479	150,893	151,307	151,721	152,135	152,549	152,963	153,377	153,791	154,205	154,619	155,033	155,447	155,861	156,275	156,689	157,103	157,517	157,931	158,345	158,759	159,173	159,587	160,001	160,415	160,829	161,243	161,657	162,071	162,485	162,899	163,313	163,727	164,141	164,555	164,969	165,383	165,797	166,211	166,625	167,039	167,453	167,867	168,281	168,695	169,109	169,523	169,937	170,351	170,765	171,179	171,593	172,007	172,421	172,835	173,249	173,663	174,077	174,491	174,905	175,319	175,733	176,147	176,561	176,975	177,389	177,803	178,217	178,631	179,045	179,459	179,873	180,287	180,701	181,115	181,529	181,943	182,357	182,771	183,185	183,599	184,013	184,427	184,841	185,255	185,669	186,083	186,497	186,911	187,325	187,739	188,153	188,567	188,981	189,395	189,809	190,223	190,637	191,051	191,465	191,879	192,293	192,707	193,121	193,535	193,949	194,363	194,777	195,191	195,605	196,019	196,433	196,847	197,261	197,675	198,089	198,503	198,917	199,331	199,745	200,159	200,573	200,987	201,401	201,815	202,229	202,643	203,057	203,471	203,885	204,299	204,713	205,127	205,541	205,955	206,369	206,783	207,197	207,611	208,025	208,439	208,853	209,267	209,681	210,095	210,509	210,923	211,337	211,751	212,165	212,579	212,993	213,407	213,821	214,235	214,649	215,063	215,477	215,891	216,30

GGCACGAGGCGGCGGGGCAGTGCGGGGATGCGCCCGGGAGGCCACAGCCTGAGGCCCTCAGGTCTCTGCAGGTTGTC
GTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCAGCAGCTTTAGCCCATGAGGAGGATGT
GACCGGGACTGAGTCAGGAGCCCTCTGGAAGCATGGGAGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCA
TCTTTCTGGCTTCGTTTGCAGCCTTGGTGTCTGGTTTGCAGGCAGCGCTACTGCCGCGCCGAGACCTGCTGCAG
GCTATGATTCTAAGCCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC'TGG
ACGATGTCTGTTATCACCAACCCCCACATTGAGGCCACTCTGGAGAATGAAGACTGGATTGCAAGATGCCCTCGGGTC
TCATGTCCCAGTGCATTGCCATCTGAAGATTTGTCACTCTGCACAGAGAAGCTTGTGGCATGACAATGGGCT
CTGGGGCCAAGATGAAGACTTCAGCCAGTGTTCAGCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGG
ATGATGTTGTGAAGTCGATGTACCCCTCCGTTGGACCCCCAAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGT
CTGTCACTCACCTGGTGTCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCCTGGACTGGATTGACCAGTCTC
TGTCTGGCTGCTGAGGAGCATTGGAAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAGCCAGATAAAGGCCCTCCCAG
GCCCTGAAGGCTTCTCTGCAGGAGCAGTCTGCAATTTAGTGCCTACAGGCCAGCAGCTAGCCATGAAGGCCCTGC
CGCCATCCCTGGATGGCTCAGCTTAGCCCTTCTACTTTTTCTTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTT
CAGCTGTGTGTCATAGTAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACTGTGGCTGGT
GAGTGGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCGCAAGAGGAGTATTGAAAAC'TGGTGGAC
TGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCTAAGAAATCAAGAGGTTTTCACAT
TAAAATTAGAATTTCTGGCCTCTCTCGATCGGTGAGAATGTGTGGCAATTCTGATCTGCATTTTTCAGAAGAGGAC
AATCAATTGAACTAAGTAGGGGTTTCTTCTTTTGGCAAGACTTGTACTCTCTCACCTGGCCTGTTTCATTTATT
TGTAATTATCTGCCTGGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGACAGGTTTGGGTTTGAAGCTGAGGAAC
TACAAAGTTGATGATTTCTTTTTTTATCTTTATGCTTGCCTGCAATTTTACCTAGCTACCACTAGGTGGATAGTAAATTT
ATACCTATGTTTTCCTTCAAAAAAAAAAAAAA

123/615

FIGURE 122

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSELELDDVVITNPHIE
AILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
DPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWIDQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

124/615

FIGURE 123

CCCTTACATCCTCCTAGGACCCGGTCGGTAGTCGTGCGCCCGAGCCCGCCGGGGGCGCAGCGCCCGAGCCGCGGCC
CTCGAGACGGGACCGAGAGCATC**ATG**GGCAGCACTGTCCCGCGCTCCGCCTCCGTGCTGCTTCTGCTGCTGCTCC
TGCGCCGGGCGGAGCAGCCCTGCGGGGCCGAGCTCACCTTCGAGCTGCCGGACAACGCCAAGCAGTGCTTCCACG
AGGAGGTGGAGCAGGGCGTGAAGTTCTCCCTGGATTACCAGGTCATCACTGGAGGCCACTACGATGTTGACTGCT
ATGTAGAGGACCCCCAGGGGAACACCATCTACAGAGAAAACGAAGAAGCAGTACGACAGCTTCACGTACCGGGCTG
AAGTCAAGGGCGTTTATCAGTTTTGCTTCAGTAATGAGTTTTCCACCTTCTCTCACAAGACCGTCTACTTTGACT
TTCAAGTGGGCGATGAGCCTCCCATTTCTCCAGACATGGGGAACAGGGTCACAGCTCTCACCCAGATGGAGTCCG
CCTGCGTGACCATCCATGAGGCTCTGAAAACGGTGATTGACTCCCAGACGCATTACCGGCTGCGGGAGGCCCAGG
ACCGGGCCCCGAGCGGAAGACCTTAATAGCCGAGTCTCTTACTGGTCTGTTGGCGAGACGATTGCCCTGTTGCTGG
TCAGCTTCAGTCAGGTGCTACTGTTGAAAAGCTTCTTCACAGAAAAACGACCCATCAGCAGGGCAGTCCACTCCT
AGCCCCGGCATCCTGCTCTAGGGCCCCCTCATGCCCCAGGCTGGAGCAGCTCTCCTAGGTCACAGCCTGCTGGGCT
GGGTCGCGTAGCCAGGGTGGAGGCAGAACGATGCTGCTGTGGTAGCCCTTTGCTTTTCATGCCCATGCTTGATT
CTTGACCTCAGCAGCTGAAGGTCTCAGAGACCAGTAATCAGAAGGCATCCGACTGCATTAAGTGTGCAGCGCTG
AAAAGACATTTACAAC TAGGCCAGGGATTAGCCACTGTGGGAGGGTGGACAGGCAATGGTT CAGTGGCCTGGCTG
TTGGCAGGAAC TCCAAGTGCCCAGGCCTCTTGGGCAGCTTAGGGCCCTGCCTCTGTTTCATGATGCATGGGTCAT
TTGTCTTGGGTGTCTTATCCCATATGGAGAAGAAAGGGGCTCTAAGTTCTGGCTCTTCTTTCTTTGGGGTTCTCT
GTACCTGAGGAAACCAGGCCCTGGGTGACTTTGCAGATCTGCTCACCCCTCGGTGAGCAACAGTGT CAGCCATGCA
AGCAGGACAGAATGGTGACTGGGTGCCCTTGGTGAGCTGTGTATTTCTTAGGAGGTAGAAAACTGTGGGAAACTG
TGGCTAATAAAAACTAAGTGTGAGCGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

105256 " 01503
20510 " 9852501

125/615

FIGURE 124

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56052
<subunit 1 of 1, 217 aa, 1 stop
<MW: 24777, pI: 5.55, NX(S/T): 0
MGSTVPRSASVLLLLLLLLRRAEQPCGAELTFELPDNAKQCFHEEVEQGKFSLDYQVITG
GHYDVDCYVEDPQNTIYRETKKQYDSFTYRAEVKGVYQFCFSNEFSTFHKTVYFDFQV
GDEPPILPDMGNRVLTALQMESACVTIHEALKTVIDSQTHYRLREAQDRARAEDLNSRV
YWSVGETIALFVVSFSQVLLLLKSFFTEKRPISRVAHS

Important features:

Signal peptide:

amino acids: 1-23

Transmembrane domain:

amino acids: 187-201

N-myristoylation sites:

amino acids: 26-32, 48-54, 131-137

Tyrosine kinase phosphorylation site:

amino acids: 82-91

Glycosyl hydrolases family 25 proteins:

amino acids: 53-61

2005-01-01 09:00:00

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTTCAGCCCCATTTGGCCCCGACGCCCTCTGT
TCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCTTAACGGA CTGCAAGATGGAGGAAGGCGGGAACTTAG
GAGGCCGTGATTAAGATGGTCCATCTACTGGTCTTGT CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCTGTCT
CAGGCTTCCTGCTTTTTCCGAAGCCTTCCCCGACATACTTCGGACTAGTGCAGAGCAAAC TCTTCCCCTCTACT
TCCACATCTCCATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCACAT
TCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCTTGAGCCTTACGCTGGCCACTGTCAACGCCCGCTGGCTGGAAC
CCCGCACCCACAGCTGCCATGTGGGGCCTGCAAACCGTGGAGAAGGAGCGAGGCCTGGGTGGGGAGGTACCAGGCA
GCCACCAGGGTCCCGATCCCTACCGCCAGCTGCGAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCT
TCCGCTACCATGGGCTGTCTCTCTTTGCAATCTGGGCTGCGTCTCTGAGCAATGGGCTCTGTCTCGCTGGCCTTG
CCCTGGAAATAAGGAGCCTCTAGCATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAAAAAAAA
AAAAAA

127/615

FIGURE 126

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

<subunit 1 of 1, 231 aa, 1 stop

<NX(S/T): 0

MEEGGN LGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISMGC AFINLCILASQ

HAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKERGLGGEVPGSHQGPDPYRQLREKDPKY

SALRQNFFRYHGLSSL CNLGCVL SNGLCLAGLALEIRSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 86-103, 60-75

Casein kinase II phosphorylation site.

amino acids 82-86

Tyrosine kinase phosphorylation site.

amino acids 144-151

N-myristoylation site.

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 54-65

G-protein coupled receptors proteins.

amino acids 44-85

2005-04-20 10:00:00

128/615

FIGURE 127

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGC
CATGATCTCCTTAACGGACACGCAGAAAAATTGGAATGGGATTAAACAGGATTTGGAGTGTTTTTCTGTTCTTTGG
AATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAAAT
TGGTTTAGAAAGAACATTAGATTCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTCTGGGTGGTGT
ATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTCAGGGG
CTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCCTGGATCCCTCCTAAATTTACCTGGAATTAG
ATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATTAACAACAAGTGAATTTGAAGACTCATTTAAAAATA
TTGTGTTATTTATAAAGTCATTTGAAGAATATTAGCACAAAAATTAAATTACATGAAATAGCTTGTAATGTTCTT
TACAGGAGTTTAAAACGTATAGCCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAAACAGGCTTCTA
CTCAAGTGAAC TAAGAAGAAGTCAGCAAGCAAAC TGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAAACTC
TTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAACTGTGGTGCCTGTTTCT
TTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCTTTTTAGAAAGTGTCCACTGCAATGGCAAAAA
TATTTCCAGTTGCACTGTATCTCTGGAAGTGATGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAC
CAAGGAAACCCCAATTTTGATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

10053586.011502

129/615

FIGURE 128

MISLTDTKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQKHMKATGFFLGGV
FVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIIRVPVLGSLNLPGIRSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

2025-10-10 10:00:00

130/615

FIGURE 129

AATTCAGATTTTAAAGCCCATTTCTGCAGTGGAAATTTTCATGAAC TAGCAAGAGGACACCATCTTCTTGATTATACA
AGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGGCCTCCTAATCCTCTGTGGTT
TTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTGATAAGTACATTTTTATCACTGGATGTGACT
CGGGCTTTGGAACTTTGGCAGCCAGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAAT
CAGGATCAACAGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGA
ATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTG
GTGTTCCCGCGTGTCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCCTATTGAAGTGAACCTGT
TTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAAGCTCAAGGGAGAGTTATTAATGTCTCCA
GTGTTGGAGGTGCGCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAAGGTTTCAATGACA
GCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAAACTTGG
CAGATCCAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATG
GAGAAGGTACATTGAAAAAGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGG
TGGTAGAGTGCATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATATATGCCGCTGGAAAAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGG
CTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGATTTCAAGAAC
ACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGATCGTGCTTATTTGGATTGCAAA
AGGGAGTCCCACCATCGCTGGTGGTATCCAGGGTCCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGT
ACATCACATAGGCAAGTCCTGCCCTGTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTT
GCCCATTCAAATGATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGT
TAAGTATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

105356 01503

131/615

FIGURE 130

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406
><subunit 1 of 1, 319 aa, 1 stop
><MW: 35227, pI: 8.97, NX(S/T): 3
MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGPHVIAACLTESGSTALKAETSE
RLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDYREPIEVNLFGLISVTNLMLP
LVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMKAFGVHVSCEPGLFKTNLADPVKVIEKKLA
IWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAA
LQDFLLLKQKAELANPKAV
```

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

132/615

FIGURE 131

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAAGCCATGAACATCA
TCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGAGTCGTTGGTGAAGTTTTTTCATTCCCTC
AGAGGAGAAAAATCTGTGGCTGGGGAGATTGTTCTCATTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTT
ATGAATTTGCAAAACGACAGAGCATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTG
AGTGCCGAAAACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCTCTC
TAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACAGTATATCCAGCCGATC
TTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAACATCCTAGGACATTTTTGGATCACAAAAG
CACTTCTTCCATCGATGATGGAGAGAAATCATGGCCACATCGTCAACAGTGGCTTCAGTGTGCGGCCACGAAGGGA
TTCCCTTACCTCATCCCATATTGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTC
AGGCCTTGGGAAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTCACCAAAAATC
CAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATAGATGGAATACTTACCAATA
AGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAGACTACAGAAGTTTCTTCCCTGAACGCGCCTCAG
CGATTTTTAAATCGTATGCAGAATATTCAATTTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATGAATAAAATA
AGCTCCAGCCAGAGATGTATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTTCACAT
TTTTTTCAGTCTTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTAATTACCTGT
CTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTTCCTTTTCATGCCTCTTAAAAACTTCTG
TGCTTACATAAAACATACTTAAAAGGTTTTCTTTAAGATATTTTATTTTTTCCATTTAAAGGTGGACAAAAGCTACC
TCCCTAAAAGTAAATACAAAGAGAACTTATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTG
TAGCCATGCCACAGAATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTT
TATCTCAACCTGGACATATTTTAAGATTGAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTTTTCATTAGCCCAA
AACGGTGCAACTCTATTTCTGGACTTTATTACTTGATTCTGTCTTCTGTATAACTCTGAAGTCCACCAAAAAGTGGA
CCCTCTATATTTCTCCTCTTTTATAGTCTTATAAGATACATTATGAAAGGTGACCGACTCTATTTTAAATCTCA
GAATTTTAAGTTCTAGCCCCATGATAACCTTTTTCTTTGTAATTTATGCTTTTCATATATCCTTGGTCCCAGAGAT
GTTTAGACAATTTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAACA
ATGGACCCAAGAGAAGAA

205110 98525001

FIGURE 132

<subunit 1 of 1, 300 aa, 1 stop

MNIILEILLLLITIIYSYLESIVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSIIVLWDINKRGVEE
TAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVNNAAGTVYPADLLSTKDEEITKTFEVNILGHFW
ITKALLPSMMERNHGHIIVTASVCGHEGIPYLI PYCSSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGF
TKNPSTRLWPVLETDEVVRSIDGILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Signal peptide:

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

CTGAGGCGGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCGCACTCGCTTTT
CCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAAGGTGAAGCCAAGAACAGCAT
TACTGATTTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGACATTAGAAAATATATTCCATGCTATCAGCT
TTTTAGCTTTTATAAATCTTCAGGCGAAGTAAATGAGCAAGCACTGAAGAAAAATATTATCAAAATGTCAAAAAGAA
TGTGGTAGGTTGGTACAATTCCGTCGTATTAGATCAGATCATGACGTTTAGAGGAGGCTCTTACAAAA
CTTGCAGGAGCATTTTTTCAAACCAAGACCTTGTTTTCTGCTATTAAACCAAGTATAATAACAGAAAGCTGCTC
TACTCATCGACTGGAACATTCCTTATATAAAACCTCAAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTTGCCAA
TCTGGGCATGCTCTGAACAACTGGGTATATAAACTGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGT
ACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCCTTAAAGGAGGTACATAAGATAAATGAAATGTTATGC
TTCTATTACAAGAGGAATTAAGAGTATATGCAAAAAAGTGAAGACAGTGAACAAGCAGTAGATAAACTAGTAAA
GGATGTAAACAGATTAAAAACGAGAAATTGAAAAAGGAGAGGAGCAGATTCAGGCAGCAAGAGAGAAGAACAT
CCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCGAAATTTCTTCA
TTTCATGTGTTATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTTAA
GCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGACAAACGATC
TAAAGCAAATAC'TGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCCAGAAACAGATGAAGAAATTTGA
AAAGATGAAGGGTTTTGGTGAATATTACGGTCTCCTACATTTT**TGAT**CTCTTTTAACTTACAAGGAGATTTTTTT
ATTTGGCTAGTGGGTAAAGCCAAACATTTCTATTGTTTTACTATGTTGAGTCTTGCAGTAGTTCATTTGTT
TTTACTATGTTTACCTGTTTGCAGTAATACACAGATAAATCTTAGTGAGTTACTTTCACAAAGTACTTTTTCAA
CATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAGACACA
TTCTTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACTTAGGGAAGACAAGTC
AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATTGAGACCATGTCTATTAATAAATAAAA
ATGGAAAAGCAAGAATAGCCTTATTTTCAAAATATGGAAGAAGAAATTTATATGAAAAATTTATCTGAGTCATTA
TTCTCCTTAAAGTGATATCTTTTGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATGCAAT
AAATTTGCAAAACATCATCTAAAATTTAAAAA

[illegible]

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLEGEVKGAEKNSITDSQMDDVEVVYITIDIQKYIPCYQLFSFYNS
SSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDLVFLLLTPSIITESCSTHRL
HSLYKPKQGLFHRVPLVVANLGMSEQLYGKTVSGSCMSTGFRAVQTHSSKFFFEEDGSLKEVHKINEMYASLQEE
LKSICKVDESEQAVDLKVDVNRLLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMS
LKNRHVSKSECNYNHHLDDVDNLLTMVEHTDPEASPASTPQIIKHKALDLDLRWQFKRKRLLDTQDKRSKANTG
SSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTE

FIGURE 135

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC
CAAGCAGCGCGCAGCGAACGCCCCGCCGCCACACCTCTGCGGTCCCCGCGGCGCTGCCACCCCTTCCCTCC
TTCCCCGCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTTCGCTGCCCCCGGAAACCCGAGGTCAACAGCC
CGCGCCTCTGCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCCCTGGCCCCGGCGCCTGGCACCGGGG
ACCGTTGCTGACGCGAGGCCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCACCAACT
CCAACCTCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCCGCTGCCGTAGCGCCGC
TTCCCCGTCCGGTCCCCAAAGGTGGGAACGCGTCCGCCCCGGCCCCGACCATGGGCACGGTTTCGGCTTGCCCCGCGCTT
CTCTGCACCCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGA
CGTCTTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATC
TGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAA
AGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTGATGAATTTCTC
AAAGAACTACTTGAAAATGCAGAGAAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAA
AATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGTGAACCTTGAAGAA
ATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAAGTCCAGTACCACCTTTACAGAT
GAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCTCGCAAATTGAAGCTC
CAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGGAGATGTCGTGAGCAAG
GTCTCCGTGGTAAACCCACAGCCCAGTGTAACCATGCCCTGTTGAAGATGATCTACTGCTCCCACTGCCGGGGT
CTCGTGAAGTGTGAAGCCATGTTTACAACTACTGCTCAAACATCATGAGAGGCTGTTTGGCCAACCAAGGGGATCTC
GATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGTCCCTTTCAACATTGAA
TCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGATAATAGTGTTCAAGTGTCT
CAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCCTCCAGCTGGACGAATTTCTCGTTCATCTCTGAAAGT
GCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACTAGTTTGGACCGA
CTGGTTACTGATGTCAAGGAGAAAAGTGAACAGGCCAAGAAAATTCTGGTCCCTCCCTTCCGAGCAACGTTTGCAAC
GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTACCTGTTT
GCAGTGACAGGAAATGGATTAGCCAACCCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGACATA
CTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCA
GAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTC
GGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATAATTCTCA
AACTCTGAGAAAAAGTGTTCATCAAAAAGTTAAAAGGCACAGTTATCACTTTTCTACCATCCTAGTGACTTTTGC
TTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAGTGCTGACTTTTGTTC
TCATTGAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCCAAACCATGTTAAACGTGGCT
AACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTTT
TTCTCATTTTCGTTTGTGGGTTTTTTTTTCCAACGTGATCTCGCCTTGTTCCTTACAAGCAAACAGGGTCCCTT
CTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATC
TTATTTAAAGAAAAAGCCCCAAAAGC

137/615

FIGURE 136

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGSTCCSQEMEKEY
SLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGHLYMQNSELFKDLFVELKRY
VVGNVNLEEMLNDFWARLLERMFRVLVNSQYHFTDEYLECVSKYTEQLKPFQGDVPRKLKLQVTRAFVAARTFAQGL
AVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMAE
RLEGPFNIESVMDPIDVKISDAIMNMQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPT
TAAGTSLDRLVTDVKEKLQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGSRYLFAVTGNGLANQGNNPEV
QVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAGKSANE
KADSAGVRPGAQAYLLTVFCILFLVMQREWR

2025-10-20 10:00:00

138/615

FIGURE 137

GCGGGCTGTTGACGGCGCTGCG**ATG**GCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTTCTCTCAGTCGGA
CTTCCTGACGCCGCCAGTGGGCGGGGCCCCCTTGGGCCGTGCGCACCACTGTAGTCATGTACCCACCGCCCGCC
GCCGCCCTCATCGGGACTTTCATCTCGGTGACGCTGAGCTTTGGCGAGAGCTATGACAAACAGCAAGAGTTGGCGGCG
GCGCTCGTGCTGGAGGAAATGGAAGCAACTGTGAGATTGCAGCGGAATATGATTCTCTTCTCCTTGCCTTTCT
GCTTTTCTGTGGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAAGCTCTGGCTTTCAGGCTAGAGGAAGA
GCAGAAAGATGAGGCCAGAAATTGCTGGGTTAAAAACCAGCAAAATCCACCCGTCTTACCAGCTCCTCAGAAGCGGA
CACCGACCCTGAGAACTTACCTGAGATTTGCTCACAGAAGACACAAAGACACATCCAGCGGGGACCACCTCACCT
GCAGATTAGACCCCCAAGCCAAGACCTGAAGGATGGGACCCAGGAGGAGGCCACAAAAAGGCAAGAAGCCCCCTGT
GGATCCCCCGCCCGGAAGGAGATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGG
CACCGAGCTCCCTTCAAGAAGAGCAGAAGTGCCACCAAGCCTCCCTGCCACCGGCCAGGACACAGGGCACACC
AGTGCATCTGAACTATCGCCAGAAGGGCGTGATTGACGCTTCTCTGCATGCATGGAAAGGATACCGCAAGTTTGC
ATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCTTTCAGTGAGTGTTTTGGCCTCGGTCTCACACTGATCGA
CGCGCTGGACACCATGTGGATCTTGGGTCTGAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTT
ACACTTTGAAAAGGACGTGGACGTCAACCTGTTTGAGAGCACGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTA
CCACCTGTCTGGGGACAGCCTCTTCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGCCTTCAGAAC
ACCATCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACCCGCCACGGTGGACCTCCGACAG
CACTGTGGCCGAGGTGACCAGCATTTCAGCTGGAGTTCCGGGAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCA
GGAGGCAGTGAGAAAGGTGACACAGCACATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGGCCATGTTTCA
CAATACCCACAGTGGCCTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTA
CCTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGAAGCCATCGAGGG
TGTCAGAACGCACCTGCTGCGGCACTCCGAGCCAGTAAGCTCACCTTTGTGGGGGAGCTTGCCACCGCCGCTT
CAGTGCCAAGATGGACCACCTGGTGTGCTTCTGCCAGGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGC
CAGCCACATGGAGCTGGCCCAGGAGCTCATGGAGACTTGTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAG
TCCCCGAGATCGTGCACTTCAACCTTTACCCCCAGCCGGGCCGTGGGACGTGGAGGTCAAGCCAGCAGACAGGCA
CAACCTGTGCGGCCAGAGACCGTGAGAGCCTGTTCTACCTGTACCGCTCACAGGGGACCGCAAATACAGGA
CTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTACACCGGTCCCTCGGGTGGCTATTCTTCCATCAACAA
TGTCAGGATCCTCAGAAGCCCCGAGCCTAGGGACAAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCT
GTTCTTGCTCTTCTCCGATGACCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCCACCTCT
GCCTATCTGGACCCCTGCC**TAG**GGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTGCTGGG
TCTGTGGCATTTTCCAAGGGCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCTCTGAACTGGCTCTGGGCT
CCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGAGGCCGTGAGTCTTGGTGTGATGCGGGGTGGG
CTGGGCCGCTGGAGCCTCCGCCTGCTTCTCCAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAG
GTCTCTGTGGGCCGACCAGAGGGGGGCTTCAGAGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGT
GCAGCTCTGCCCGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGGTCTGGAGGGGCTGCCGTGACTCCAGAG
GCCTGAGGCTCCAGGGCTGGCTCTGGTGTTTACAAGCTGGACTCAGGGATCCTCCTGGCCGCCCCGAGGGGGCT
TGGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCTCCAGTGAATGGGTCTTTTCGGTGGAGATAAAAG
TTGATTTGCTCTAACCGCAA

2025-10-29 10:59:00

139/615

FIGURE 138

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529
><subunit 1 of 1, 699 aa, 1 stop
><MW: 79553, pI: 7.83, NX(S/T): 0
MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSFGESYDNSKSWRRRSCWRKW
KQLSRLQRNMILFLLAFLLFCGLLFYINLADHWKALAFRLSEEQKMRPEIAGLKPANPPVLPAPQKADTDPENLP
EISSQKTQRHIQRGPPHLQIRPPSQDLKDGTOEEATKRQEAPVDRPEGDPQRTVISWRGAVIEPEQGTELPSSR
AEVPTKPPLPPPARTQGTPVHLNRYRQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWI
LGLRKEFEEARKWVSKKLHFEKDVDVNLFESTITILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKIPYS
DVNIGTGVAHPPRWTSDSTVAEVTSIQLEFRELSRLTGDKKFQEAVEKVTQHIIHGLSGKKDGLVPMFINTHSGLF
THLGVFTL GARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRTHLLRHSEPSKLTFFV GELAHGRFSAKMDHL
VCFLPGTLALGVYHGLPASHMELAQELMETCYQMNRMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLRPET
VESLFYLYRVGTGRKYQDWGWEILLQSFSRFRTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDD
PNLLSLDAYVFNTAEHPLPIWTPA
```

Important features of the protein:

Transmembrane domain:

amino acids 21-40 and 84-105 (type II)

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

140/615

FIGURE 139

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGACCTGAGTCATC
CCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACTTACAGCTGCACCGACAGTTG
CGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGA
ATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAAT
GTGAGTGCAAAGATTGGTTCCCTGAGAGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGT
GCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAAAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATT
CCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTGAGCGC
CCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTC
ACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTGTTGT
TGCTCTCTCTAGTGTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTAC
CTGAAAGATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAAC
AGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

2025-09-01 10:00:00

141/615

FIGURE 140

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFMTVSGLPKKQC
PCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

2025-09-01 10:00:00

142/615

FIGURE 141

AATGGCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCTTGACCTGCTATGCAGACGA
CAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCAAAGCCAGACTTCCCCAAATTCCTAAG
CCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGAGTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTAT
GGAATTTGATGATAATGAAGGAAAACATTCATCAAAGTGAATCCTCAGGACACACCCATGTGGCTCCTGGACAA
TCCAAGAGCAGCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAACTCC
CAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACCCTGACTGCATTTTGC
TTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTGATGGAGAGGAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

20510" 92525001

143/615

FIGURE 142

MAVLVLRRLTVVLGLLVLFILTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVEFILRSMRSTGFM
EFDDNEGKHSSK

20570" 98625001

144/615

FIGURE 143

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAGAGCTGGTCTG
CC**ATGG**ACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCCCCCTGCACCTCATGGCTCTGC
TGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCCCCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCA
AGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTAC
TGGAGCTGGGCTGCGGAACCGGAGCCAACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAA
ATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGG
CTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTGCTCTG
TGCAGAGCCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGC
ATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAACACATTGGGG
ATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAATGGAACGAC
AGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCT
CCAAGGCACTCATTGTCTCCTTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGA
GAGGGACCT**TAG**CAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGG
ACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTCCCTTTCCTT
CGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGTCTCTAGGAACTGGTCACAAAAG
TCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGCA
CCAGGGAGAATCAGAGATGCTGGGGATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTCTCAATATTTTT
TAATAAATAGACGAAACCAG

20510"9265001

145/615

FIGURE 144

MDILVPLLQLLVLLLTLPHLMLALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGLTGASGKVALL
ELGCGTGANFQFYPPGCRVTCDDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDMRQLADGSMDVVVCTLVLCVS
QSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQVFEPTWKHIGDGCCLTRETWKDLNAQFSEIQMERQ
PPPLKWLPVGP HIMGKAVKQSFPSKALICSFPSLQLEQATHQPIYLP LRG T

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
1990	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022

146/615

FIGURE 145

GTGGGATTATTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAGTTGCCTCATCGCAGGCAGATGTTGGGGC
TTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAAACTAATATTTATATGACAGAAGAAAA
AGATGTCATTCCGTAAAGTAAACATCATCATCTTGGTCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATA
ACTTCCTCAGCTTGAGCAGTTTGTAAAGGAATGAGGTTACAGATTTCAGGAATTGTAGGGCCTCAACCTATAGACT
TTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTGAAG
ACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTTCAGCACAACTCGCTCCAATGTGATTTTCTACATTG
TTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAA
TTGTCAATTTTGACCCCTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTT
TAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAA
TTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTTCAGAAGATT
GTGATTTCAGCCTCTACTAAAGTTGTTCATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATA
AAAAGGAAAGAATTTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAACC
TGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGT
ATAGCAGAACCCTGGCTGGTAGCATCACAAACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCG
ATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTACCTCAGTTTGTAAAGGCTGCCA
AGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGAAAAATGGT
ATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAACAGAA
TTTGAAGTGTAAAGCAAGCATTCTCAGGAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTC
AATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGT
CAGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTG
AATGACTGGAAAGAAGAACTGATATGGCTAGTTTCAGCTAGCTGGTACAGATAATTCAAACCTGCTGTTGGTTTTA
ATTTTGTAACTGTGGCCTGATCTGTAAATAAACTTACATTTTTC

20510.9525001

147/615

FIGURE 146

MSFRKVNIIILVLAVALFLLVLHNNFSLSSLLRNEVTD SGIVGPQPIDFVPNALRHAVDGRQEEIPVVIAASED
RLGGAIAAINSIQHNTRSNIIFYIVTLNNTADHLRSWLNSDSLKSIRYKIVNFDPKLLEGKVKEDPDQGESMKPL
TFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNTALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYK
KERIRKLSMKASTCSFNPGVFVANLTEWKRQNI TNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTID
PMWNVRLGSSAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

MSFRKVNIIILVLAVALFLLVLHNNFSLSSLLRNEVTD SGIVGPQPIDFVPNALRHAVDGRQEEIPVVIAASED

FIGURE 147

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTTCCTCCAAGCA
AGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCTATTTGCATCTGTTTTGATAA
ATGATGTTGACACCCTCCACCGAATTCTAAGTGGAAATCATGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCC
TCGCATTAGCCTTGTCTTTGGCCATGATGTTTACCTTCAGATTTCATCACCACCCTTCTGGTTTCACATTTTCATTT
CATTGGTTATTTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCA
GCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGG
CAGTGCTGCTCGTCTTGATTTTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCACAAATA
AAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCATTTTTCTTCTGGGTCC
TCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGAATATA
AGCCCCTTTCGGGCATTTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCATCCTTG
CGTGCCAGCAAATGACTATAGCTGGGGCAGTGTTACTTGTATTTC AACAGAAGTAAAAATGATCCTCCTGATC
ATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCT
CTGTGGTGAGGATTCCGAGAATCATTGTCTGTACATGC AAAACGCAC TGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATG
CATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAATCTTGTCCAAGA
ACTCAAGTCACTTTACATCTATTAACCTGCTTTGGAGACTTCATAATTTTTCTAGGAAAGGTGTTAGTGGTGTGTT
TCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGG
TAGCTTTTTTTGCCTACTTAGTAGCCCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGT
GTTTTGCTGTTGATCTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTT
TCGTAAAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGAGGGAA
CAGAACTCCAGGCCATTGTGAGATAGATAACCCATTTAGGTATCTGTACCTGGAAAACATTTCTTCTAAGAGCCA
TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTTTTTAAAGACCTAATAAACCCCTATTC
TTCTCAAAA

149/615

FIGURE 148

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYYDYTNDSLIELDTERENMKC
VLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFAILIFFWVLWVAVLLSLGTAG
AAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIAGAVVTCYFNRSKNDDPPDHPILSSLSILFFY
HQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFC
TSAKDAFKILSKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVAHSFL
SVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

2025-10-27 10:00:00

[illegible]

MRTVVLTMKASVIEMLVLLVLTGVHSHNKETAKKIKRPKFTVPQINCVDKAGKIIDPEFIVKCPAGCQDPKYHVYGT
TDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWRESFIVLESKPKKGVITYPSAI
TYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMLLAVTVAVATPTTLPRPSPSAASTTISIPRPQSVGHSRQE
MDLWSTATYRTSSQNRPRADPGIQRQDPSGAAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDG
STSIGKRFRFIQKQLLADVAQALDIPGAPLMGVVQYGDNPATFHLKTHNTSRDLKTAIEKTIQRGGLSNVGRA
ISFVTKNFFSKANGNRSGAPNVVVVMDGWPTDKVEEASRLARESGINIFFITIEGAANEKQYVVEPNFANKAV
CRTNGFYLSLHVQSWFGLHKTLQPLVKRVCDDTLRLACSKTCLNSADIGFVIDGSSSVGTGNFTVLQFVTNLTKFE
EISDTDTRIGAVQYTYEQRLFQFDKYSSKPDILNAIKRVGYWSSGGTSTGAAINFALQLFKKSKPNKRKLMILI
TDGRSYDDVRI PAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARDHSFFVDFDNLHQYVPRIIQNICTEFNSQPRN

FIGURE 151

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGCGCTGAGAGGACACGAGCTCT**ATG**CCTTTCCGGCT
GCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGTGCGCCAGGTCCCGACGGCTCCGCGCC
AGATCCCGCCCACTACAGTTTTTCTCTGACTCTAATTGATGCACCTGGACACCTTGCTGATTTTGGGGAATGTCTC
AGAATTCCAAAGAGTGGTTGAAGTGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGA
AACAAACATTCGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGAGGC
TGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCGAAAACCTCTCCAGCCTTTCAGAC
CCCCACTGGCATGCCATATGGAACAGTGAACCTTACTTCATGGCGTGAACCCAGGAGAGACCCCTGTACCTGTAC
GGCAGGGATTGGGACCTTCATTGTTGAATTTGCCACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGT
GGCCAGAGTGGCTTTGATGCGCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCT
CACTGGCAAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGTGAAAGG
AGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTTCTAGAGTATAACAAAGCCATCCGGAACCTACACCCG
CTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACTGTGTCCATGCCAGTCTTCCAGTCTTGGAGGC
CTACTGGCCTGGTCTTCAGAGCCTCATTGGAGACATTGACAATGCCATGAGGACCTTCTCAACTACTACACTGT
ATGGAAGCAGTTTGGGGGGCTCCCGGAATTCACAAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTA
CCCACTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACTCTCTAGAACT
CGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGATTTGCAACAATCAAAGATCTGCG
AGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCTGCGCGAGACTGTGAAATACCTCTACCTCCTGTTTGA
CCCAACCAACTTCATCCACAACAATGGGTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGG
GGCTGGGGGGTACATCTTCAACACAGAAGCTCACCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAA
GGAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCCTCTCTCAAACGGAGCAGGTGAAATTTTCAGAA
AAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACACTCTTCTCACCAGAAAACCATGACCA
GGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCACTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTT
GGCATTACTGGGACAGGTTTTCTAGACTCCTCATA**TA**ACCACTGGATAATTTTTTTATTTTTTATTTTTTTGAGGCT
AAACTATAATAAATTGCTTTTGGCTATCATAAAA

[illegible]

FIGURE 153

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAAAGTGAAGAAAACA
GAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCTTTCGT
CACTCACCTGTTCTTGCCCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTTAACTGGATGAACATCACCCACG
CCTATTCCCAAGGGCCACCAAGACTGAATTTGGATACAGTGTCTTACAACATGTTTGGGGGTGGACAGCGATGGA
GCTGTTGGGGCCCCCTGGGATGGGCCCTTCAGGCGACCGGAGGGGGACGTTTATCGCTGCCCTGTAGGGGGGGC
CCACAATGCCCATGTGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTCATCTCTGCTGTGAATAT
GCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTTCATGGTGAGCTAAAGGAGAGGGTGGTGGCAG
TGTCTCTGAAGGTCCATAAAAAGAAAAAGAGAAGTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGT
TAAAAACCTTAGAAAGCAAAAGGTAGGTAATGTACGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTT
TGAGGGTGCCCTCCCAAGCCTGGGAGTAACATATTTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCTCGTGCT
TGTGGCAGCTCTGCTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCAGCCTCAGGGAAGCCTGGCA
CCCACTGGCCCAACGTGAGCCAGAGAAGGCTGAGTACTTGGTTCACCAAGGAGATCTGGGTGGGAAAAAGATG
GGGCAAAAGCGGTATGATGCCTGGCAAAAGGCCTGCATGGCTATCCTCATGGTACTACCTAATGTGCTTGCAAAAGCT
CCATGTTTCTTAACAGATTTCAGACTCCTGGCCAGGTGTGGTGGCCCCACACCTGTAATTCTAGCACTTTGGGAGGC
CAAGGTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAAATCCATCTCTACTAA
AAAAAAAAAAAAATACAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTCATCTACTCGGGAGGCTAAGACA
GGAGACTCTCACTTCAACCAGGAGGTGGAGGTTGCGGTGAGCCAAGATGTGTGCCCTCTGCACTCTAGCGTGGGTG
ACAGAGTAAGCGAGACTCCATCTCAAAATAATAATAATAATCAGACTCCTTACAGGAGTCCATGATCTG
GCCTGGCACAGTAACATCATGCCTGTAATCCCAACATTTGGGAGGCCAACGACGAGGAGTGTCTGAGGTCTGGA
GGTTTGAGACCAGCCTGGGCAACATAGAAAGACCCATCTCTAAATAAATGTTTTAAAAAT

155/615

FIGURE 154

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSFPNLDHHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPWDGPGSDRRGDVYR
CPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGFMVS
```

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

2025-09-24 10:00:00

156/615

FIGURE 155

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGAGGCTCCCAGCTGC
AGCGTCCCCGCCCCGCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCCCTCGGGGACCAACAAGCCTGGCAGG
GTCTCACTTTTGTGGCCAGGCTGGAGTTCAAGTCCATGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAA
GCGATCCTGCTGAGTAGCTGGGACTACAGGACAAAATTAGAAGATCAAAATGGAAAAATATGCTGCTTTGGTTGAT
ATTTTTTCACCCCTGGGTGGACCCCTCATTTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAAGGT
ACCCCGGATTGTCAAGTGAAGGACTTTCCATCTCACCAGCCCCGATTTGAGGCAGATGCTAAGATGATGGTAAA
TACAGTGTGTGGCATCGAATGCCAGAAAGAAGTCCCAACTCCAGCCTTTCTGAATTGGAGGATTATCTTTCCCTA
TGAGACTGTCTTTGAGAATGGCACCCTGAACCTTAACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCA
AAATATCACCACAAAGGGAGTATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCAGCAGCAGGTTGAGCATCTT
GGACAAAAGGTTCTTAACCAATTTCCCTTTGAGCAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCATTCTCAT
TTCCCTCAGCATGTTCTAACTGCTGCCCCTGTGTTTCAATGATGGAAAGGACTATGTCAAAGGGAGTAAAAAGCT
AAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAGAAACGTCGAGGTTCTAAGAGGAGCAGGAGAGA
AGCTAGTGGTGGTGACCAAGAGAGGGGTACCAAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAA
AAAATCTGGCCGGGGTCAAGAGATTGCCGAAGGGAGGCTTCTTTTCAAGTGGACCCGGGTCAAGAATACCCACAT
TCCGAAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGAGCTGAAGCG
TGCTCAGAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAATGCCTGGTGGAAATGATCCACTT
CTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTATCGGTTTTGTCAGTGTGTCCGACGAATCCAATGATCT
CCTTTACCAATACTGCGATGCTGAGTCGGGCTCCACCGGTTCCGGGGTCTATCTGCGTCTGAAAGATCCAGACAA
AAAGAATTGGAAGCGCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTTCAAGAGGA
CTACAACGTTGCTGTTGCGATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTACCGGGAACGATGCCAA
TTGTGCTTACGGCTTAAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAATCACAGAGAAAACCAGCTCTGCTTA
CCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTTGAACCTCTGTCAATAGCATTTCAACATTTTTTCAAATCA
GGAGATTTTCGTCCATTTAAAAAATGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATAT
ACTCTTCTTTACATGGTGTGATGAGTTTCATTTGTAGAAAAATTTTGTTCCTTCTTAAAAATTAGACACACTTTAA
ACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCTACTCTAAGAAGAAT
CTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAAGGTAGATGGTAAAGCAATTAGTATCAGA
ATAGAGACAGAAAGTTACAACACAGTTTGTACTACTCTGAGATGGATCCATTGAGCTCATGCCCTCAATGTTTAT
ATTGTGTTATCTGTTGGGTCTGGGACATTTAGTTTATGTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGC
ATTATAAACAAAACATAAATGTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAATGGGA
GAAATAGTTTGTCTATGAAATAAACCTAGTTTGAAGATAGGGAAGCTGAGACATTTTAAAGATCTCAAGTTTTTA
TTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGGAAGACACTTCACAAATTATGAATGATCATGT
GTTGAAAGCCACATTATTTTATGCTATACATTCTATGTATGAGGTGCTACATTTTATAGGACAAAGAATTCTGTAA
TCTTTTTTCAAGAAAGAGTCTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTTCTCTG
ATTAGTAATTTTATGATATGTCCTTTCTTAAAAATGAATAAAATTTATGAATATGA

157/615

FIGURE 156

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVCGIECQKELPTPSL
SELEDYLSYETVVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQVYGTDSRFSILDKRFLTNFPFSTAVKL
STGCSGILISPQHVLTAACHCVHDGKDYVKGSKKLRVGLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQE
RAKGRRRKKSGRGQRIAEGRPSFQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIK
KMPGGMIHFSGFDNDRADQLVYRFCSVDESNDLLYQYCD AESGSTGSGVYLRLKDPDKKNWKRKIIAVYSGHQW
VDVHGVQKDYNAVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 157

GGGACCCATGCGGCCGCTGACCCCCGGCTCCCTAGAGGCCAGCGCAGCCGAGCGGACAAAGGAGCATGTCCGCG
CCGGGGGAAGGCCCGTCTCTCCGGCCGCCATAAGGCTCCGGTCGCCGCTGGGCCCGCGCCGCGCTCCTGCCCGCCCG
GGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCGCTCGCCCCGAGGCCCGCGCCGAGCATGGAGCCACCC
GGACGCCGCGGGGGCCGCGCGCAGCGCCGCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGA
GGCGCGCGCGCGCGCGCGCGCGCTGCCCGCGCTGCAAGCAGCATGGCGCGGCCCGAGGGGCTGGCAGG
CGCGCGGGCGCCCGCAGGGCAAGGTGGTGTGCAGCAGCCCTGGAACCTCGCGCAGGTCCTGCCCCAGATACCTCTG
CCCAACCGCACGGTCAACCCTGATTCTGAGTAACAATAAGATATCCGAGCTGAAGAATGGCTCATTTTTCTGGGTTT
AGTCTCCTTGAAAGATTGGACCTCCGAAACAATCTTATTAGTAGTATAGATCCAGGTGCCCTTCTGGGGACTGTCA
TCTCTAAAAAGATTGGATCTGACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTTCGAGGACTCACCAAT
CTGGTTTCGGCTAAACCTTTTCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACCTTTTGATTATCTTCGCTCATTA
CGTCTTTTGGAAATTCAGACTGAGTATCTTTTGTGTGACTGTAACTATCTGAGTACTGTGGATGCATCGCTGGGTAAAGGAG
AAGAACATCACGGTACGGGATACCAGGTGTGTTTTATCCTTAAGTCACTGTCAGGCCCAACCAGTCACAGGCGTGAAG
CAGGAGCTGTTGACATGCGACCCCTCCGCTTGAATTGCCGCTCTTTCTACATGACTCCATCTCATCGCCAAGTTGTG
TTTGAAGGAGACAGCCTTCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTGGTATCAG
GATGGGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCACAACCTGCTCCTTG
ATTGCAAGTGCCCTAACCATTTCTAATATTCAAGCTGGATCTACTGGAATTTGGGGCTGTGATGTCAGACAAA
CGTGGGAATAATACGAGCATGTGATATTGTGGTATTAGAGAGTTCTGCACAGTACTGTCTCCAGAGAGGGTG
GTAAACAACAAGGTGACTTCAGATGGCCCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAAC
ACCCATGGCAGTGGGATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGC
TTTTTGGGCAGATGATGATTATTCTCGCTGTGAGTATGCAATGATGTCAC'TAGAGTTCTTTATATGTTTAATCAG
ATGCCCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGCTTACACTGTGGAAGCGCCAACCTTT
CTCGCAAAATGGATGTTATATTTGTGGCAGAAATGATTGAAAAATTTGGAAGATTTACCAAGGAGGAAAAATCA
AAAGAGCTAGGTGACGTGATGGTTGTGACATTGCAAGTAACATCATGTTGGCTGATGTAACGTGTCTGTGGCTGGCG
CAGAGGGAAGCTAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTGGGA
GCTCACGTTTATTCAACATATTCACCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCTACTGGCTTCACGGGG
ATGACCTGTACCGTGTTCAGAAAAGTGGCAGCCTCTGATCGTACAGGACTTTTCGGATTATGGGAGGCGGGATCCA
GAGGGAACCTTGATAAGCAGCTGAGCTTTAAGTGCATGTGTTCAAATACATTTTCGAGTCTGGCACTAAAGGTA
TGTTACATCTCGCAATCATTTTAAGACTATTTACAGTTAAATAGAATGCTCCAAATGTTCTGTGTTGCAAAATAA
CCTTATTAAAAAGATTTTTTTTTTGTGCAGGAAGATAGGTATTATTGCTTTTTGCTACTGTTTTAAAAAACTAAACCAG
GAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCCTTTGATTCCCTTTCTTCACATAAAAAATATCA
GAAATTACATTTTTATAACTGCAGTGGTATAAATGCAAAATATACTATTGTTACATGTGAAAAAATTTTTATTGACT
TAAAGGTTTATTATTGTTTTTTTTTGCTCCTGATTTTAAAGACAATAAGATGTTTTTCATGGGCCCTTAAAGTATC
ATGAGCCTTTGCACTGCGCCTGCCAAGCCTAGTGGAGAAGTCAACCTTGAGACCAGGTGTTTAATCAAGCAAGC
TGATATATCAAAATTTTTTGGCAGAAAAACAATAATGTCATATATCTTTTTTTAAAAAAAGTATTTTCATTGAAGCA
AGCAAAATGAAAGCATTTTTTACTGATTTTTTAAATTTGGTGCTTTAGATATATTTGACTACACTGTATTGAAGCAA
ATAGAGGAGGCACAAC'TCCAGCACCC'AAATGGAACCACATTTTTTTTCACTTAGCTTTTCTGTGGGCATGTGTAATT
GTATCTCTCTGCGGTTTTTTTAACTCTCACAGTACTTTTATTTTCTGTCTTGTCCTCAATAATATCACAAACAATATTC
AGTCATTTTAATGGCTGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCACTCAATAAA
TATTGAATGAATGAACGAAAAAATAAAAAAAAAA

159/615

FIGURE 158

MEPPGRRRRGRAQPPLLLPLSLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGKVVCSSLELAQVLP
PDTLPNRTVTLILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGAFWGLSSLKRLDLTNNRIGCLNADIFR
GLTNLVRLNLSGNLFSSLSQGTFDYLASLSLEFQTEYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPV
TGVKQELLTCDPPELPSFYMTTPSHRQVVFEGDSLFPQCMASYIDQDMQVLWYQDGRIVETDESQGIFVEKNMIH
NCSLIASALTISNIQAGSTGNWGCCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDWRWPRTLGITAYLQ
CTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNLNTNAVATARQLLAYTVE
AANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLADERVLWLAQREAKACSRIVQCLQRIATYR
LAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCTVFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSL
ALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336, 433-437,
453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57, 57-63, 99-105,
123-129, 142-148, 162-168, 317-323, 320-326, 384-390, 403-409, 554-560

160/615

FIGURE 159

GGGGAATCTGCAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTGGCTGCTGTTGTTT
CTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGAAAGTATTTATTGACCAAATTAACAGG
TCTTTGGAGAATTACGAACCATGTTCAAGTCAAACTGCAGCTGCTACCATGGTGTCTATAGAAGAGGATCTAACT
CCTTTCCGAGGAGGCATCTCCAGGAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATC
ACTAAGAACAGACTGTACCGGGAAAATGACTGCATGTTCCCTCAAGGTGTAGTGGTGTGAGCACTTTATTTTG
GAAGTGTATCGGGCGTCTCCCTGACATGGAGATGGTGTCAATGTACGAGATTATCCTCAGGTTCTTAAATGGATG
GAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACCATGATATCATGTATCCTGCTTGGACATTT
TGGGAAGGGGGACCTGCTGTTTGGCCAATTTATCCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTG
GTAAGGTGACGAGCACAGTGGCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCA
GAACGAGATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAACTTGTGATGCAGAATACACCAAAAACAGGCC
TGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGTGGATCACTGCAAATACAAG
TATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTAAACACCTCTTCCTGTGTGGCTCACTTGTTTTC
CATGTTGGTGATGAGTGGCTAGAATTCTTCTATCCACAGCTGAAGCCATGGGTTCACTATATCCAGTCAAAACA
GATCTCTCCAATGTCCAAGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGG
GGAAGCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTGAGTGAATAC
TCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTATTCCCAAATGTTGAAAACCTGAA
CTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCAACAGATCTCAGATATCCTACGGTGAGAAGCTTACC
ATAAGCTTGGCTCCTATACCTTGAATATCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAG
AGCAACTCTTGAGAAAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGG
ACCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTTCAATTTCTTAAGACCAATCACAGCTTGTGCC
TCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCATGTGATGATGCCCTTTGTCCCATTA
TTTGGAGCAGAAAATTCGTCAATTTGGAAGTAGTACAACCTATTGCTGGAATTGTGAAATTATTCAAGGCGTGATC
TCTGTCACTTTATTTTAATGTAGGAAACCTATGGGGTTTATGAAAATACTTGGGGATCATTCTCTGAATGGTC
TAAGGAAGCGGTAGCCATGCCATGCAATGATGTAGGAGTTCTTTTTGTAAAACCATAAACTCTGTTACTCAGGA
GGTTTCTATAATGCCACATAGAAAAGAGCCAATTGCATGAGTAATTATTGCAATTGGATTTTCAAGTTCCCTTTTT
GTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

20510 9555001

161/615

FIGURE 160

MEWWASSPLRLWLLLFLLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCCSSQNCSCYHGVIEEDLTPFRGGISRK
MMAEVVRRKLGTHYQITKNRLYREND CMFPSRCSGVEHFILEVIGRLPDMEMVINVRDYPQVPKWMEPAIPVFSF
SKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDLFREDLVRSAQWPWKKKNSTAYFRGSRTSPERDPLILLS
RKNPKLVDAEYTKNQAWKSMKDTLGKPAAKDVHLVDHCKYKYLNFNFRGVAASFRFKHLFLCGSLVFHVGDWLEF
FYPQLKPWVHYIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWENLLSEYSKFLSYNVT
RRKGYDQIIPKMLKTEL

20510" 2225001

FIGURE 161

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGCCAAA
CCATGTCTTTTTCTGTTTTTCAGAGTAGTTTCAACAACAGATCTGAGTGTTTTAATTAAGCATGGAATACAGAAAA
CAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTT
GGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCG
CGTGCTTCTGAGCTGCTGTGG**ATG**GCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCAGTACCTCAGCCTTCCCCACTACAATGTG
AAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCTTCCCCACTACAATGTG
ATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTACAGACAAGACTTTCACTTCACACTTCGA
GAGCATTCAAACTGCTCTCATCAAAATCCATTTCTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCC
AGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTA
GGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATA
ATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGCATTCAAGTGGGTAAGTGAAG
TTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTAT
CTTTTAAACCTAAACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTT
TACCAAAAAACCCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTAT
ATAATGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAGAACCCATCAAGTTTGAAGATGTT
TATGTCGGGATCTGTTTGAATTTATTAAGAGTGAACATTCAATTTCCAGAAGACACAAATCTTTTCTTTCTATAT
AGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACT
TTTTGGCAGGTCATGCTAAGGAACACCACATGCCATTAT**TAA**CTTCACATTCTACAAAAAGCCTAGAAGGACAGG
ATACCTTGTGGAAAGTGTTAAATAAAGTAGGTAAGTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACT
GAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGGCCCTTCAA
GATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGTAAAGAAATTAATAGGACCAACAATTTG
GACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAA
AACAATGTAGAGTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAAT
TTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTTTG
GTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTTTAAATTTACTTCAACTTTGTGTT
TTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAGTGAATCATTCTTTACATGCAACATTTTCCAGT
TACTTAACTGATCAGTTTATTATTGATACATCACTCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTA
ATCTCTTGGACTTTGTAAATATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

163/615

FIGURE 162

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHFTLREHSNCSSH
QNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLALSLEDEHLLYGDIIRQDFLD
TYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHIS
YQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVC
QLRRVIAAHGFSSKEIITFWQVMLRNTTCHY

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 163

CATTTCTGAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATATTAACTTTTTAGG
AGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCGCTCATATATAGGAAAATCGCATATGG
TCCTAGTATTAAATTCCTATTGCTTACTGATTTTTTTGAGTTAAGAGTTGTTATATGCTAGAAATATGAGGATGTG
AATATAAATAAGAGAAGAAAAAGAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAACTGGTT
TGTTTACATGCAAGCTTATAGTTGAAATATTTTTTCAGGAATTAC**ATGA**ATGACAGTCTTCGAACCAATGTGTTTG
TTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTTCAGATTCCGTTGCCAA
CTCGTCCCCATTGGTTTTCTTCTTTTTTGGTACTACAGAAGAGGAAATCCAGGAAATCTGCATAGAAACACTTAGGC
TTTATACCAGAAAAAGCCAACTATGAATTACTGGAAAAAGAAGTAGAAAAAGAAAAGTAGCCTTACAAGAAG
CCAAATTTAAAGCAAAGGGATTGAATCCGGATGGAACCTCCAGCCCTTTCAACCCCTGGGTGGATTTTTCTCCAGCCT
CCAAGCCATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAGTCAAAA
AAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAAGACAGCAAGAGAAGTAGAA
ATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACGTTCTAGATCACATACTCCAAGAAGACACTATA
ATAATAGGCGGAGTCGATCTGGAACATACAGCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTC
GAAGACATCATAATCATGGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGAC
ATGGTCATAAAAGGAAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAGCCAAGAAAC
ACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTTGAGAGGTCCCATAAAAGCAAGC
ACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCGCT**TGA**CTTTCTCTTCCTTTGAGCCTGCATCAGTTCT
TGGTTTTGCCTATCTACAGTGTGATGTATGGACTCAATCAAAAACATTAAACGCAAACCTGATTAGGATTTGATTT
CTTGAAACCCCTCTAGGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACAT
TAAAATGCCCTAGCAGTATCTAATTAATAAACCATGGTCAGGTTCAATTGTACTTTATTATAGTTGTGTATTGTTT
ATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTTATACAGATAAAATTGCAGACACTGT
TCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAATACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATT
TTTACAAGGAAATAAAATACAAATCTTGTTTTTTCTAAAAAAAAAAAAAAAAAAGT

165/615

FIGURE 164

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLFLFGTTEEEIQEICIETLRLYTRKKPNYELLEKE
VEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEKSPISINVKTVKKEPEDRQQASKSPYN
GVRKDSKRSRNSRSASRSRSTRSRSRSHTPRRHYNNRRSRSGTYSSRSRSRSHSESPPRRHHNHGSPHLKAKH
TRDDLKSSNRHGHKRKKSRSRSQSKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSRSGHGRHR

165/615

GGTTCCTACATCCTCTCAICTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTTATTAACGTGGCTTA
ATCTGAAGGTTCTCAGTCAAATTCCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCCTAAAGGAGCTT
GGCTGGTTTTGGGCCCTTGTAGCTGACAGAAGGTTGGCCAGGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGC
TTCTGTTGCTGTTCTTGCCCTGGCTCAGTCTGCTTAACATTGCAATGTGGGCAACCTGCACCTTCTGTATT
CAGAACTCTGTAAAGGTGCCTCCCACTACGGCTGACCAAGATAGGAAGAGCGCTCACAAGATGGTGTCCAG
ACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCGAGAGTTTCTGCAGCTGCCACCATTCTCTTAATGACAG
ACGAGCCTGGCCTAGACAACCCCTGCCTACGTGTCTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACT
CTGGCCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCAITTTAAAAAATAA
ATCGAGCTTTGAGTGTCTTTCGAAGGACAAAGAGCGGGAGTGCATTGCCAAACCATGCCGACAGGGCCAGGGAAA
ATTTCTGAAAACCACTAGCCCCGTGAAGTCTTTCCAAGTTGTACCCACTGATTCCAGATGGTGAATATACAGCA
TCAAGATCAATCGAGTAGATCCCAAGTGAAGCCCTCTATTAGGCTGGTGGGAGGTAGCCAAACCCCACTGGTTC
ATATCATTATCCAACACATTTATCGTGTATGGGTGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC
TAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCTGCGGCAGCCCTGCCAGG
TGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCCGGATGCCTACAGAC
CCCCGATGACAGCTTTTCATGTATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAAATAAACTGGTGC
AGGTGAGTAGGCTGGGGTTTTGATCTTTCAATGTGCTGGATGGCGGTGCGCATATCGACATGGTTCAGCTTGAGG
AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAGTGCGGCTCATCTGATT
AGGCCAGTGAAGACGTGTTCACTCGTGTGTCGCCAGGTTTCGGCAGCGGAGCCCTGACATCTTTTCAGGAAG
CCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACTCCCAAGCCCTCCATCCTA
CAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGTGAATCTCTCGGCATGACCGTCGAGGGG
GAGCATCACATAGAGAATGGGATTTGCCATCTATGTGTCATCAGTGTGAGCCCGGAGGAGTCAAGCAGAGATG
GAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCTGAAGTACAGAGGTCAGCCGGAGTGAGGCAG
TGCCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGTTTGGAAAGTCAAAGAGTATGAGCCCGAGGAAG
ACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCTGGGTCA
TGTGGCTGGAATTACACGGTGTCTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTCTGG
GCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAA
TGATACATCTGCTTGTGGCAAGACTGCTGAAAGAACTTAAAGGAAGTACTACTTAACATTTGTTCTTGGCCTG
GCACTTTTTTTATAGAACTCAATGATGGGTGAGAGGAAAAACAGAAAAATCACAATAGGCTTAAGAAGTTGAAACACT
ATATTTATCTTGTGATTTTTATATTTAAAGAAAAGAAATACATTGTAAAAATGTGAGGAAAAAGTATGATCATCTAA
TGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTTAAACTACTAGTTTTTTTTTCAGTGTGGAGGAT
TTCTCATTACTCTACAACATTTGTTTATATTTTTTCTATTCAATAAAAAGCCCTAAAACAACATAAATGATTGATT
TGTATACCCCACTGAATCTCAAGCTGATTTTAAATTTTAAATTTTGGTATATGCTGAAGTCTGCCAAGGGTACATTAT
GGCCATTTTTTAATTTACAGCTAAAATATTTTTTAAATGCAATTGCTGAGAAACGTTGCTTTCATCAAACAAGAA
AAATATTTTTTCAGAAGTTAAA

167/615

FIGURE 166

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPSPEVSAAATIS
LMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINRALSVLRRTKSGSAVANHADQ
GRESENTTapeVFPrlyHLIPDGEITSIKINRVDPSESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPG
DIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIK
LVRKVDEPGVFI FNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDI
FQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVI
SRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPPSDWSP
SWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRS
TSGMHAACLARLLKELKGRITLTIVSWPGTFL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

168/615

FIGURE 167

GGGAAAGCCATTTTCGAAAACCCATCTATACAAACTATATATTTTCATTTCTGCTGCTAGCTGCCTTGGGCCTCAC
AATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGAAATGGAGTTGATCCCAACCATAACATCGTGGAG
GGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTGGTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAAATCA
TGAACCTCTGGCTGTTGATCAAAAGAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAAGAAGCT
AGCAGAAGACTCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACATCAA
CGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAACCTCAAATTGGGAGGCCAACCCACAGAACAGCATT
CTGGGCCAGGCTGTAAATCAGAAATTGTCGTCGTACATGCTCAACAGCATTGCTTTTTTTCCCCAAAATTAACACATT
GTGGAGAAGTGATGATACTCTCCCCCTTACCTTTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATT
AAACCTTGCAGCAAGGGACCTTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATT
CCTGTATCATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

2024-10-23 10:00:00

169/615

FIGURE 168

MELIPTITSWRVLLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLAEDSTWPPINRTDY
SGDGKNGFYINGGYESHEQIPKRKLKLGQPTEQHFWARL

corrected

FIGURE 169

CGCTCGGGCACCAGCCGCGGCAAGG**ATG**GAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCT
TCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATGTG
TCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGAAAGAGGGGAAGTCGTGGGTATACCATCCC
TTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAC TGCAA
GAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCG
AGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGGA
AAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTTCATGCTAAACCTGGGTTTGTATCCAATAAGATTTGT
CATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTTCGTGATGGAGACAACCGCGATGG
CCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACTCCACGTCCT
CTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTTCATGCCATTTATGAGGAGATCACAGCATGCTCCTCATC
CCCTTGTTCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTGGCAGGCTATAC
TGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCCTGGGGGGCCAGTCAATGGGTACCAGAA
AATAACAGGGGGCCCCGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCCTGGTGTCTTTCTTTTGTAACAA
CTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCCATCTG
CATAAAAGCCTGCCGAGAACCAAAGATTTTCAACCTGGTGAGAAGGAGAGTTCTTCCGATGCAGGTTTCACTCAAG
GGAGACACCATTACACCAGCTATACTCAGCGGCCCTTCAAGCAGAAACTGCAGAGTGCCCCCTACCAAGAAGCC
AGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACC
CTTCTACCGCCCGCTTGGGCAGCAGCAGGAGCATGTCTGAGGACTGGGAAGTGGAGTGGGCGGGCACCATCCTG
CATCCCTATCTGCGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGCGCTGGCCGTGGCAGGCAGC
CATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGGTTCTAGTCTGCGAGCGGTGC
CCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAAGGTCACCATGATCAAGAC
AGCAGACCTGAAAGTTGTTTTGGGGAAATTTCTACCGGGATGATGACCGGGATGAGAAGACCATCCAGAGCCTACA
GATTTCTGCTATCATTCTGCATCCCACTATGACCCCATCCTGCTTGATGCTGACATCGCCATCCTGAAGCTCCT
AGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAGTCCGGGATCTCAGCACTTCCTTCCA
GGAGTCCCATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACT
GCGCTCTGGGGTGGTCACTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCCACTGAG
TGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGG
CATCGCGGCTGTGTCTTCCCGGACGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAG
CTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCCTTCAACCAAGGTGCTGCCTTTTAAAGACTGGATTGAAAG
AAATATGAAAT**TGA**ACCATGCTCATGCACTCCTTGAGAAGTGTTTCTGTATATCCGTCTGTACGTGTGTATTGCG
TGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAACCTGGCTGTGCCAGGGCTTCTGACTTCAGGGACAAA
ACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCGTCCACTACTAGGACAGCCAATTGGAA
GATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACCATATACAAAACCTCTCCACTCCACTGACCTGGT
GGTCTTCCCAACTTTTCACTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAGGCCCTTT
TGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
GTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCCATCTCTTGTACACATTTTAAATAAAATAAGGGTTG
GCTTCTGAAC TACAAAAA
AA

171/615

FIGURE 170

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYTIPCCRNEENE
CDSCLIHPGCTIFENCKSCRNGSWGGLTDDFYVKGFYCAECRAGWYGGDCMRGQVLRAPKGQILLESYPLNAHC
EWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKN
FDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLI
NGRHAKIGTVVSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLHQLY
SAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIE
NITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADLKVVLG
KFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLDDKARISTRVQPICLAASRDLSFSQESHITVAG
WNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGI PVSVDNMFCASWEPTAPSDICTAETGGIAAVSFPG
RASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

171/615

172/615

FIGURE 171

CTGTCGTCTTTGCTTCAGCCGAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTCCAAGTGTGGCTTA
ATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCTCTGCTAAGACCGCTGCCATGCCAGTGACGGTAACCCG
CACCACCATCACAACCACCACGACGTATCTTCGGGGCTGGGGTCCCCCATGATCGTGGGGTCCCCTCGGGGCCCT
GACACAGCCCCCTGGGTCTCCTTCGCCTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGT
GGGCGCCTGGACGGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGCTTCTGCTTCTCCGTGACCCCTGATCAT
CCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCCATCACCTTCGCCTG
CTATGCGGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCACCACCTATGTCCAGTTCCTGTCCCACGGCCG
TTCGCGGGGACCACGCCATCGCCGCCACCTTCTTCTCCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTG
GACCCGGGGCCCGGCCGGCGAGATCACTGGCTATATGGCCACCGTACCCGGGGCTGCTGAAGGTGCTGGAGACCTT
CGTTGCCTGCATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGGCCCTGGAGTGGTGCCT
GGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGGGAGTGCACCAACGTGCT
ACCCATCCCCCTTCCCCAGCTTCTGTGCGGGCTGGCCCTTGTCTGTCTGCTCCTCTATGCCACCGCCCTTGTCTCT
CTGGCCCCCTTACCAGTTCGATGAGAAGTATGGCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAG
CCATGCCTACTACGTGTGTGCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTA
TGTGGCTGACCTGGTGCATCTGCCCCACCTGGTTTTTGTCAAGGTCTTAAGACTCTCCCAAGAGGCTCCCGTTCCC
TCTCCAACCTCTTTGTTCTTCTTGGCCGAGTTTTCTTTATGGAGTACTTCTTTCCTCCGCCTTCTCTGTCTTTC
CTCTTCCTGTCTCCCTCCCTCCACCTTTTTCTTTCCTTCCCAATTCCTTGCACTCTAACCAGTTCCTTGGATGC
ATCTTCTTCCCTCCCTTCTCTTGTCTTTCCTTCTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
AGCTGTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT
GAGTGCAGTGGTGCATCTCAGCTCACTGCAACCCCGCCTCCTGGGTTCAAGCGATTCTCCTCCCCCAGCCTCC
CAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTTTCTCACTCTTCTTTTCTCTATC
TCTTTTCTGGGTTGCCGTGTCGGCTTCTTATCTGCCTGTTTTGCAAGCACCTTCTCCTGTGTCTTGGGAGCCCT
GAGACTTCTTTCTCTCCTTGCCCTCCACCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGACGCCGTCC
ATGCCACAGCCCCCAAGGGGCCCATTGCCAAAGCATGCCTGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC
GTGTGTGTGTGTGTGTGTGTGGGGGTGGGGGTGGGTAGCTGGGGATTGGGCCCTCTTCTCCCAGTGGAGGAA
GGTGTGCAGTGTACTTCCCCTTTAAATTAAAAAACATATATATATATATATATATTTGGAGGTCAGTAATTTCCAATGG
GCGGGAGGCATTAAGCACCGACCCCTGGGTCCCTAGGCCCGCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAG
AATTTTTGCCAGGCTTACAGAACCCCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTCAT
CCCAACTATTCTCTGTGGTATGAAAAAG

173/615

FIGURE 172

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727

<subunit 1 of 1, 322 aa, 1 stop

<MW: 35274, pI: 8.57, NX(S/T): 1

MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTGSMGNWSMFTWCFC
FSVTLIILIVELCGLQARFPLSWRNFPITFACYAALFCLSASIIYPTTYVQFLSHGRSRDHAIATFFSCIACVA
YATEVAWTRARPGEITGYMATVPGLLKVLETFVACIIFAFISDPNLYQHQPALWCWVAVYAICFILAAIAILLNL
GECTNVLPPIPFPSFLSGLALLSVLLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRLAVAILT
AINLLAYVADLVHSAHLVFKV

Important features:

Transmembrane domains:

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192, 205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

20250401 14:00:00

[illegible]

175/615

FIGURE 174

MFLATLSFLLPFAHPFGTVSCEYMLGSPPLSSLAQVNLSPFSSHVKVHMDPNYCHPSTSLHLCSLAWSFTRLLHPPL
SPGISQVVKDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALSAFSSYSDDLSEGEQEARFAAGVAEQFAIAEAK
LRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLGPHLQDLFTGHRFSRPVRQGSVEPESDCSQTVPDTLCSS
LCSLEDGLLGSPARLASQLLGDELLLAKLPPSRESAFRSLGPLEAQDSLNSPLTESCLSPAEEEPAPCKDCQPL
CPPLTGSWERQRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:
amino acids 1-15

Casein kinase II phosphorylation site.
amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232, 285-289,
324-328

Tyrosine kinase phosphorylation site.
amino acids 44-52

N-myristoylation site.
amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 11-22

20250909 09:00:00

FIGURE 175

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTTCCCTGGCACCCCTCCTGC
TCAGTGCAGACATTGTACACTTAACCCATCTGTTTTCTCTAATGCACGACAGATTCCTTTCAGACAGGACAACCTG
TGATATTTTCAGTTCCTGATTGTAAATACCTCCTAAGCCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTT
CTTCATCTGCAAAATGGGCATAATAACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACC
AATACCAAAGAAGCCTACAAATGTTGGCCTTAGCCAAAAATTCTGTTGATTTCACCGTTGTTTTATTCACTTCTATC
GGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTTTTAAACAATGGAAAA
TAAACCTATTTCTTTGGAAAGTGAAGCAAACCTAAACTCAGATAAAGAAAAATATAACCACCTCAAATCTCAAGGC
GAGTCATTCCCCCTCCTTTGAATCTACCCAACAACAGCCACGGAATAACAGATTTCTCCAGTAACCTCATCAGCAGA
GCATTCTTTGGGCAGTCTAAAACCCACATCTACCATTTCACACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAA
AGTGCCTTTGGAAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACCTGCTCT
GTCTTCAGAAAACCTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACTCCTGATAACAGTTCCATTACAGTTAG
CATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTGATAGTGGAACCAAGTGGATGGCTTACCACAAA
CAGTGATAGCTTCACTGGGTTTACCCCTTATCAAGAAAAACAACCTCTACAGCCTACCTTAAAAATCACCAATAA
TTCAAACTCTTTCCAAATACGTGAGATCCCCAAAAAGAAAAATAGAAATACAGGAATAGTATTCGGGGCCATTTT
AGGTGCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGGAAAAAGGAAAACGGATTTCATT
TTCCCATCGGCGACTTTATGACGACAGAAATGAACCAGTTCTGCGATTAGACAATGCACCGGAACCTTATGATGT
GAGTTTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAATGATTAGCCATGCCAGAAAGTGAAGAAAATGCACG
TGATGGCATTCCATGGATGACATACCTCCACTTCGTACTTCTGTATTAGAACTAACAGCAAAAAGGCGTTAAACA
GCAAGTGTCTATCTACATCCTAGCCTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTCACGTGGA
TTTTGTCAAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCATCCAAAGGTTTTCT
TTTCTTACAATTTTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATTTGTATTTTAGTAGTATTTTCTTAGT
AGAAAATATTTGTGGAATCAGATAAACTAAAAGATTTACCATTACAGCCCTGCCTCATAACTAAATAATAAAAA
ATTATTCACCAAAAAAATTCTAAAACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATT
CAAGATTGCATTTTCTTAAATGAAAATTGAAAGGGTGCTTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAGGA
CATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTCTAGTACGTTATAATT
TTCTAGATCAGCACACACATGATCAGCCCACTGAGTTATGAAGCTGACAATGACTGCATTCAACGGGGCCATGGC
AGGAAAGCTGACCCTACCCAGGAAAGTAATAGCTTCTTTAAAAGTCTTCAAAGGTTTTTGGGAATTTTAACTTGTC
TTAATATATCTTAGGCTTCAATTATTTGGGTGCCTTAAAACTCAATGAGAATCATGGT

177/615

FIGURE 176

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNIAEVFKTMENKPISLESEANLNSDKENITTSNLKASHSPPL
NLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNAPIADEDLLPISAHPNATPALSSSENFT
WSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVEPSGWLTTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFNP
TSDPQKENRNTGIVFGAILGAILGVSLTLVGILLCGKRKTDSEFSHRRLYDDRNEPVLRLDNAPEPYDVSEFGNSS
YYNPTLNSAMPESEENARDGIPMDDIPPLRTSV

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 235-262

N-glycosylation site.

amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159, 163-167, 218-222,
225-229, 298-302, 307-311

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ACCAGGCATTGTATCTTCAGTTGTTCATCAAGTTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTTCTTGCCT
GCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACTTTCCACCTTTCCCTA
CAAATTCGATTACTGTTGCTGTTGACTTTTGTGCCTGACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGG
TGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTTCATGGCTAATTTCCATAAGACCCCTCATTTTGGGGGAAGGGAAA
AATCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGTTGACAACTGCTCCTTGCTGTCTCCTTACCTCAGAG
CCAGAGCAAGCTCATTTTCAAACAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCGAGAGG
CCGGTATCGCCCTCAGGAATGTAAAGCTTTTACAGAGGGTGCCTCCTCGTTCCCAACCGGAAACAGAGAGAAAACA
CCTGATGTACCTGCTGGAACATCTGCATCCCTTCTGTCAGAGGCAGCAGCTGGATTATGGCATCTACGTCTATCCA
CCAGGCTGAAGGTAAAAAGTTTAAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAA
TTGGGACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAGCA
TCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGGTGTTAC
TGCCCTTAAGCAGAGAGCAGTTTCTCAAGGTGAATGAAATTTCTTAACAACTACTGGGATGGGGAGGCGAAGACGA
TGACCTCAGACTCAGGGTTGAGCTTCAAGAAATGAAATTTCCCGGCCCTGCTCGAAGTGGGTAAATATACAAAT
GGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCCACGAGT
CTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATATATCAACAT
CACAGTGGATTTCTGGTTTGGTGCATGACCCTGGATCTTTTGGTGATGTTTGGGAAGAACTGATTCTTTGTTTGCA
ATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTT
TCCTTTTGTATTTTCTTAGCAGAGCTCCTGGTGATGTAGAGTATAAAAAGCTTGTAACAAAGACAGCTTTCTTAG
TCATTTTGTATCATGAGGGTTAAATATTGTAATATGGATACCTTGAAGGACTTTATATAAAAGGATGACTCAAAGGA
TAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGATAAAAA
GGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAAAGGTACGAAGA
TACAATACTGTTATTCAATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGTCAGGTGAGAAGGCGTCCACAAAA
GAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCCGC
TGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCAGGTATGCCTTC
CAGTGAATGCCACAGAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGG
ATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAAATATGTCTATCAAATACCTCTGTAGTAAAT
GTGAAAAAGCAAAA

179/615

FIGURE 178

MGFNLTfHLSYKfRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTILGKGKTLTNEASTKKVELD
NCPSVSPYLRGQSKLI fKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPHRNREKHLMYLLEHLHPFLQR
QQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFI fHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYR
LRYSGYFGGVITALSREQFFKVNGFSNNYWGWGGEDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAE
RMKLLHQVSRVWRDGLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

2025-04-04 10:00:00

180/615

FIGURE 179

CGTGGGCCGGGGTTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGAGTTCTCGAGCTCCAGCTGC
ATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCAATGGCCCCAGGCAGTGTGGTTCGCGCCTCGG
CCGCATCCTCTGGCTTGCCCTGCCTCCTGCCCTGGGCCCCGGCAGGGGTGGCCCGAGGCCCTGTATGAACCTCAATCT
CACCACCGATAGCCCTGCCACCACGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAG
CCTGGCCCTGCCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAAGAT
GGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTTCGGCCACGTGCCCGGGGAATTCCCGGTCTCTGTCTGGGTAC
TGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTGGTCCTCCCCATCACAGAGTTCTCGTGGG
GGACCTTGTGTGTCACCCAGAACACTTCCCTACCCTGGCCAGCTCCTATCTCACTAAGACCGTCCCTGAAAAGTCTC
CTTCTCCTCCACGACCCGAGCAACTTCCCTCAAGACCGCCTTGTCTTCTCTACAGCTGGGACTTCGGGGACGGGAC
CCAGATGGTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCTCAAAGT
GGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGACCGGGGACTTCTCCGCCTC
GCTGAAGCTGCAGGAAACCTTCGAGGCATCCAAGTGTGGGGCCCCACCCTAATTCAGACCTTCCAAAAGATGAC
CGTGACCTTGAACCTTCTGGGGAGCCCTCCTCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCCTCCCGCTGGA
GGAAGGGGAGTGCCACCCCTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCCTGGGGA
CTACTGCTTCAGCATCCGGGCGGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGTGTGGCCCTC
CAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGATGTTGGCCTTCATCATGTACAT
GACCCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAGAACCCGGAGCCACCCTCTGGGGTCAAGGTGCTGCTG
CCAGATGTGCTGTGGGCCCTTCTTGCTGGGAGACTCCATCTGAGTACCTGGAAATTGTTCTGTGAGAACCACGGGCT
GCTCCCGCCCCCTCTATAAGTCTGTCAAACTTACACCGTGTGAGCACTCCCCCTCCCCACCCCATCTCAGTGTTA
ACTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAATTTGCGTGGGGCTG
TTGGCCTGGATCATCCATCCATCTGTACAGTTAGCCACTGCCACAAGCCCCCTCCCTCTCTGTACCCCTGACCC
CAGCCATTACCCCATCTGTACAGTCCAGCCACTGACATAAGCCCCACTCGGTTACCACCCCTTGACCCCTTACC
TTTGAAGAGGCTTCGTGCAGGACTTTGATGCTTGGGGTGTTCCTGTGACTCCTAGGTGGGCCTGGCTGCCCCAC
TGCCCATTCCTCTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCTCCCCCAGACAGCCCTACCT
GTGCCAGAGAGCTAGAAAGAAGGTCTAAAGGGTTAAAAATCCATAACTAAAGGTTGTACACATAGATGGGCACA
CTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACACACACACACAGAAATATAAACACATG
CGTCACATGGGCATTTAGATGATCAGCTCTGTATCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCCTAGAGC
TGAAAGGAAATTTGACCTCCAAGCAGCCCTGACAGGTTCTGGGCCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCA
GTTCTTGCGCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCTGGATGGGGGGCAGGACTAATAC
TGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTGAAACTTTCACTGAGGAAA
AGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCAGCACTTTGGG
AGGCCGAGGCGGGTGGATCACGAGATCAGGAGATCGAGACCACCTGGCTAACACGGTGAAACCCCGTCTCTACT
AAAAAATACAAAAGTTAGCCGGGCGTGGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGA
GAATGGTGCAGAACCCGGGAGGCGGAGCTTGCAGTGAGCCAGATGGCGCCACTGCACTCCAGCCTGAGTGACAGA
GCGAGACTCTGTCTCCA

CGTGGGCCGGGGTTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGAGTTCTCGAGCTCCAGCTGC

181/615

FIGURE 180

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLALPADAHLYRFHWI
HTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGFVVLPITEFLVGDVLTQNTSLPWSS
YLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDTQMVTEDSVVYYNYSIIIGTFTVKLKVVAEWEEVEPDATRAV
KQKTGDFSASLKLQETLRGIQVLGPTLIQTFQKMTVTNLNFGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYN
LTHTFRDPGDYCFSSIRAENIIISKTHQYHKIQVWPSRIQPAVFAPCATLITVMLAFIMYMTLRNATQQKDMVENP
EPPSGVRCCCQCCGPFLLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 181

CGGACGCGTGGGCGGCGGGCTGCGGAACTCCCGTGGAGGGGCGCGTGGGCCCCTCGGGCCTGACAGATGGCAGTGGC
CACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCGGCCCCGCCGGTTCGTGGGGCCCAGGGT
CCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCATGCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGG
CCTGGGCGCGCCACGGCCGCGCGAGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCG
CGCCGAGGAGGCGGCGGGTCACTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCCAGAGCCTGGCGTCAG
CGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCGCCTTCTGCCAGGAAAT
GCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGGATCTTCCAGTGCCCTTACATGAAGACTGA
AGATGGGTTTGAGATGCAGTTCGGAGTGAACCATCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCT
CAAAAGTTCAGCTCCCAGCAGGATTGTGGTAGTTTCTTCCAACTTTATAAAATACGGAGACATCAATTTTGATGA
CTTGAACAGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAACCTGGCTAACATTTCTTTTACCAG
GGAAGTACCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTGGTATTGTACGGACAAATCT
GGGGAGGCACATACACATTTCACTGTTGGTCAAACCACTCTTCAATTTGGTGTGATGGGCTTTTTTCAAACTCC
AGTAGAAGGTGCCCGACTTCCATTTATTTGGCCTCTTCACTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGG
GGATTGTAAAGAGGAAGAACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACTCTGGGATATCAGTGA
AGTGATGGTTGGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAGTTATATCTG
TGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTGATATTGGAATAGCCTGCTAAGA
GGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTTGGGATAAGAGAATTTTCAAGCAAAAGATGTTTAAAT
ATATATAGTAAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAATAATATAACTGGGCAAGCA
TGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGT
TTCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTGTTGTGTGGAATTATCTGCCTGGTGTGTGCA
CACAAGTCTTACTTGAATAAAATTTACTGGTAC

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 182

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
```

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPVRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAELLRLGARVIMGCR
DRARAEAAAGQLRRELROAAECGPEPGVSGVGELIVRELDIASLRSVRAFCQEMLQEEPRLDVLINNAGIFQCPY
MKTEDGFEMQFGVNHLGHFLLTNLLGLLKSAPSRIVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANI
LFTRELARRLEGVNTVTVNLHPGIVRTNLGRHIHILLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSG
RYFGDCKEEELLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 183

AACAGGATCTCCTCTTGCACTCTGCAGCCCAGGACGCTGATTCAGCAGCGCCCTTACCGCGCAGCCCCGAAGATTC
ACT**ATGGT**GAAAATCGCCCTTCAATACCCCTACCGCCGTGCAAAAAGGAGGAGGCGCGGCAAGACGTGGAGGCCCTC
CTGAGCCGCACGGTCAGAACTCAGATACTGACCGGCAAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCC
TCTGGGAGATGTATGCTTACTCTCTTAGGCCTTTTCATTTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATT
TACAAGTACTTCATGCCCCAAGAGCACCATTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCAAAT
TCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGAGGATGACAACATTGCA
ATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCTGCAGCAATTATTCATGACTTTGAAAAGGGAATG
ACTGCTTACCTGGACTTGTTGCTGGGGAAGTCTATCTGATGCCCTCAATACTTCTATTGTTATGCCTCCAAAA
AATCTGGTAGAGCTCTTTGGCAAAGTGGCGAGTGGCAGATATCTGCCTCAAACCTTATGTGGTTCGAGAAGACCTA
GTTGCTGTGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGAAAGTCC
TTCCGCCTTCGTGCGCAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAAATGCTGGAAGATTAGACAC
TTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**TAA**GAGGCAACAGATAGAGTGTCCTTGGTAATA
AGAAGTCAGAGATTTACAATATGACTTTAACATTAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTAC
TCTATTGCTTATGCTTTAAAAAAGGAAAAAAAAAAAAAACTACTAACCACCTGCAAGCTCTTGTCAAATTTTAGTT
TAATTGGCATTGCTTGTTTTTTGAACTGAAATTACATGAGTTTCATTTTTTTCTTTGCATTTATAGGGTTTAGAT
TTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATCCATCCGTTGTTTTTTTTGTTTGTGTGT
TTTTTCTTTTCCTTTAAGTAAGCTCTTTATTTCATCTTATGGTGGAGCAATTTTAAAATTTGAAATATTTTAAATT
GTTTTTGAACTTTTTGTGTAAAATATATCAGATCTCAACATTGTTGGTTTCTTTTGTTTTTTCATTTTGTACAAC
TTCTTGAATTTAGAAATTACATCTTTGCAGTTCCTGTTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAAT
TTTCATGAGACAGTCATTTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGGAATGCACAAAAT
TGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAACCCTATAATAAATTTTACTCTATAC
AAAAAAAAAAAAAAAAAAAA

185/615

FIGURE 184

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSFILAGLIVGGACIY
KYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIAIIDVPVPSFSDSDPAAIHDFEKGMT
AYLDLLLGNCYLMLPLNTSIVMPPKNLVELFGKLASGRYLPQTYVVREDLVAVEEIRDVSNLGIFTYQLCNRKSF
RLRRDLLLLGFNKRAIDKCWKIRHFPNEFIVETKICQE

Type II transmembrane domain:

amino acids 53-75

N-glycosylation site.

amino acids 166-170

Casein kinase II phosphorylation site.

amino acids 35-39, 132-136, 134-138

N-myristoylation site.

amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 63-74

accession number

186/615

FIGURE 185

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATGG**CACTGAGCTCCCAGATCTGGGCCGCTTGCCT
CCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCACAACAGACGGGACAACCTGCAGAGCT
GCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCATGTTCCAGAGGCCAAGGAGGCGAGACACCCA
CTTCCCCTATCTGCATTTTCTGCTGCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAGA**ACCT
ACCTGCCCTGCCCCCGTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGG
CTGGTTCTTTTGT'TTCCAA
AA

186/615

187/615

FIGURE 186

MALSSQIWAACLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPICIFCCGCCHR
SKCGMCCKT

protein sequence

FIGURE 187

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGCCTGGATCTTCCAC
CATGTTCCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGCATCTCCCTGACTGTCTCTTACCCT
 CCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTGGAGTCTCCTTTGGTATCCGCAAACCTACATGAAAAGTCT
 GTTAAAAATCTTTGCGTGGGCTACCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCC
 CTACACCAACGGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCTGTCGAAGTGGTAG
 TAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCCGGAAGGAATGGAGAC
 CATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTGGAGTCTTGGAACTGCTGAGCAGAACC
 TTATAACTTCCAGTACATCAGCCTTCGGCTCACGGTCCTGTGGGGGTAGGAGTGCTGATTTCGGTACTGCTTTCT
 GCTGCCGCTCAGGATAGCACTGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTT
 GCCAAATGGGAGGTTTAAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCGAGCGCT
 GACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCTGTGTGGCCAATCATACCTC
 ACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCCATGGTGGGTCAAGTGCACGGGGGACTCATGGG
 TGTGATTTCAGAGAGCCATGGTGAAGGCTGCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCT
 GGTGGCTAAGAGACTGACTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCAGAAAGGAACCTG
 CATCAATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTGAATTTGGAGCCACAGTTTACCCTGTTGCTAT
 CAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGATGGTGACGTACCTGCTGCGAAT
 GATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTGCCTCCCATGACTAGAGAGGCAGATGAAGATGCTGT
 CCAGTTTGCGAATAGGGTGAAATCTGCCATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCC
 GAAGAGGGAGAAGGTGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCA
 CAAGGACAGGAGCCGCTCT**TGAG**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAACGGGCTCAGAGC
 TGGAGTTGCCGCCGCCGCCCTTCTGCTGTGTCTTTCCAGACTCCAGGGCTCCCCGGGCTGCTCTGGATCCCAG
 GACTCCGGCTTTTCGCCGAGCCGAGCGGGATCCCTGTGCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGAT
 GCTGCTGGGTGTTGCGACCCAGGACGAGATGCCCTGTTTCTTTTACAATAAGTCTGTTGGAGGAATGCCATTAAAG
 TGAATCCCCACCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGATGG
 CGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCGCTCTCCAGGAAAGGC
 ACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCAGCCTTGGAGCTCTGCAGACATGATAGGAAG
 GAACTGTCTATCTGCAGGGGCTTTTCAGCAAAATGAAGGGTTAGATTTTTATGCTGCTGCTGATGGGGTTACTAAA
 GGGAGGGGAAGAGGCCAGGTGGGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTTCGTACTCCAGGCTAACCTG
 AACTCCCCATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGGGGAATG
 GTGGTGATTCTTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGAAGGACACATCACGTTTCA
 GTTTTCAAGTACAGGCCCAAAAACGGGGCACGGCAGGCCTGAGCTCAGAGCTGCTGCACTGGGCTTTGGATTTG
 TTCTTGTGAGTAAATAAAACTGGCTGGTGAATGA

189/615

FIGURE 188

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSGIRKLYMKSLLKIFAWATLRMERGAKEKNHQLYKP
YTNGIIAKDPTSLEEEIKEIRRSKGSSKALDNTPEFELSDIFYFCRKGMETIMDDEVTKRFSAAEELESWNLLSRTN
YNFQYISLRLTVLWGLGVLIRYCFLPLRLIALAFTGISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRAL
TAIITYHDRENRPNGGICVANHTSPIDVILASDGYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHL
VAKRLTEHVQDKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYLLRM
MTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTFKEEQQKLYSKMIVGNH
KDRSRS

189/615

190/615

FIGURE 189

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCACGTCTCCTCCAGG
GATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCACACCTGGCAGGCCCAGGCTGTTCCAC
CATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGACGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAA
GGCAGCCCCCTGCTAAAGGAGGAAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCCAGGAGAC
CTGGGAGGACAAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGTCTA
CACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAGGCTCCCGGGAGCTCTA
CATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGGGCCCTGCAGCTGCTGCGAGGCAGTGGGGG
CTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGAGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGC'TGGGGGA
CTCTGTCCGCTTGGGCCAGTTTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCG
GGGCTGTGTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTGGAAGAC
TCTGCTCTTGGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCTGAAGTCCAACATCTGCCACTTAGGAGC
CCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAGCAGCCTTGAGAAGCAAGAACATGGTTCCGGAC
CCAGCCCTAGCAGCCTTCTCCCAACCAGGATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTA
TGTGATGGGGACTTCTGGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTGCAATGTGGA
GACATGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

2025-10-20 10:00:00

191/615

FIGURE 190

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHHALLRESWEAAQET
WEDKRRGLTLPPGFKAQNGIAIMVYTNSNTLYWELNQAVRTGGGSRELYMRHFPPKALHFYLIRALQLLRGSGG
CSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSSLDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKT
LLLAPGEFQLSGVGP

2025.03.03

192/615

FIGURE 191

GTGGCTTCATTTAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCCCTCATCTATAT
CCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGAC
TTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCAT
ACAGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTA
CTCCCTGAAGCTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCA
GCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCA
GAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATAC
CTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGA
AAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCGTGCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAA
GCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCT
CAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAA
GAGAGTGGACATTTGTGCGGAAACTCCTAACATATGCCCCATTCTGGAGAGAAACACAGAGTACGACACAATCCC
TCACACTAATAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGC
AGTGCACTCCCCTAAGTCTCTGCTCA

192/615

193/615

FIGURE 192

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQPEGGTIIVTQN
RNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSKPKVTMGLQSNKNGTCVTNLT
CCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFICVARNPVS RNFS SPILARKLCEGAADDPDSS
MVL LCLLLVPLLLSLFVLGLFLWFLKRERQEEYIEKKRVDICRETPNICPHSGENT EYDTIPHTNRTILKEDPA
NTVYSTVEIPKKMENPHSLLTMPDTPRLFAYENVI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 193

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCCAGAGCAGCCCCGGGCACCAGCACGGACTCTCTCTTCCAGCCCAGG
 TGCCCCCACTCTCGCTCCATTTCGGCGGGAGCACCCAGTCTGTACGCCAAGGAACTGGTCTTGGGGGCACC**ATG**
 GTTTTCGGCGGCAGCCCCCAGCCTCCTCATCCTTCTGTTGCTGCTGCTGGGGTCTGTGCCCTGCTACCGACGCCCGC
 TCTGTGCCCCCTGAAGGCCACGTTTCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCC
 CCGAGCCTCCCGCCACCCTGGACCCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGGGGC
 CCATCACCCCCCACCAACTTCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGATGCTGATTGCTGTGGTG
 GGCTCCCTGGCCTTTCTGCTGATGTTTCATCGTCTGTGCCGCGGTTCATCACCCGGCAGAAGCAGAAGGCCCTCGGCC
 TATTACCCATCGTCCTTCCCCAAGAAGAAGTACGTGGACCAGAGTGACCGGGCCGGGGGCCCCCGGGCCTTCAGT
 GAGGTCCCCGACAGAGCCCCGACAGCAGGCCCCGAGGAAGCCCTGGATTCTTCCCGGCAGCTCCAGGCCGACATC
 TTGGCCGCCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATGGTGGAG
 GGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCAGGGACATGGGGTCCCAGTG
 GAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCTTGGAGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGG
 GAGCTGGAAGGGTCTCTCTTTAGCCCAAGAACCCAGGGACAGTGGGTCCCCCGAAAGCCCCCTGTGCTTGC
 AGCAGTGTCCACCCAGTGTCT**TAA**CAGTCTTCCCGGGCTGCCAGCCCTGACTGTGCGGGCCCCAAGTGGTCACCT
 CCCCCTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCTGACACTCCCTCCTTGGCCTCCCTGTGGTGCCAATCC
 CAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGAGGAATCTTACCAAGTGCCATCA
 TCCTTCACCTCAGCAGCCCCAAGGGCTACATCCTACAGCACAGTCCCCTGACAAAGTGAGGGAGGGCACGTGT
 CCCTGTGACAGCCAGGATAAAACATCCCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCCTGCCCGGCCCAAAC
 TACTTTTAAAAACAGCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAATACTGCTCTTAATTTTCTGAAG
 GTGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTAAATCCTCTCAAGCGCTCTC
 CAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGGGATCAGGTTGAATGAATGGAACCTCT
 TCCTGTCTGGCCTCCAAAGCAGCCTAGAAGCTGAGGGGCTGTGTTTGGGGGACCTCCACCCCTGGGGAAGTCCGA
 GGGGCTGGGGAAGGGTTTCTGACGCCAGCTGGAGCAGGGGGGCCCTGGCCACCCCTGTTGCTCACACATTGT
 CTGGCAGCCTGTGTCCACAATATTCGTCACTCCTCGACAGGGAGCCTGGGCTCCGTCCTGCTTTAGGGAGGCTCT
 GGCAGGAGGTCTCTCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCTCCAGGTGCTGAGATA
 TAATGCACCAGCACATAAACCTTTATTCCGGCCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAGA

195/615

FIGURE 194

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852
><subunit 1 of 1, 283 aa, 1 stop
><MW: 29191, pI: 4.52, NX(S/T): 0
MVSAAAPSLILLILLLLLSVPATDARSVPLKATFLEDVAGSGEAEAGSSASSPSLPPPWTPALSPTSMPGPQPTTLG
GPSPTNFLDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKASAYYPSSFPPKKKYVDQSDRAGGPRAF
SEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSPTRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVP
VETPEAQEEPCSGVLEGAVVAGEGQGELEGSLLLAQEAQGPFVGPPEPCACSSVHPSV
```

Signal peptide:
amino acids 1-25

Transmembrane domain:
amino acids 94-118

N-myristoylation site.
amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199, 211-217, 238-244,
242-248

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

GAAAGACGTGGTCCCTGACAGACAGACAATCCCTATTCCCTACCAAATGAAGATGCTGCTGCTGCTGCTGTGTTTGGGA
CTGACCCTAGTCTGTGTCCATGCGAAGAAGCTAGTTCTACGGGAAGGAACCTTTAATGTAGAAAAGATTAAATGGG
GAATGGCATACTATTATCCCTGGCCTCTGACAAAAGAGAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTG
GAGCAAATCCATGTCTTGGAGAATCCCTTAGTTCTTTAAAGTCCCATCTGTAAGAGATGAAGAGTGCTCCGAATTA
TCTATGGTTGCTGACAAAAGAAAAGGCTGGTGAATATTCGTGACGTATGATGGATTCAATACATTACTATA
CCTAAGACGACTATGATAACCTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGAAACCTTCCAGCTGATG
GGGCTCTATGGCCGAGAACCAAGATTTGAGTTTCAGACATCAAGGAAAAGTTTGCACAACTATGTGAGGAGCATGGA
ATCCTTAGAGAAAATATCATTGACCTATCCAATGCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCTTGA
GCCTCCAGTGTTGAGTGGACACTTCTCACCAGGACTCCACCATCATCCCTTCCCTATCCATACAGCATCCCCAGTA
TAAATTCGTGATCTGCATTTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACCTAGGAT
ACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAATTTTTTCATGAAATTATTCCT
CTTCCTGTTCAATAAATGATTACCTTTGCACCTTAA

197/615

FIGURE 196

MKMLLLLCLGLTLVCVHAAEEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQIHVLENSLVLKVH
TVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLINEKDGETFQLMGLYGREPDLSDDIKE
RFAQLCEEHGILRENIIDLSNANRCLQARE

2025-10-20 09:28:00

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 197

GGCTCGAGCGTTTCTGAGCCAGGGGTGACC**ATG**ACCTGCTGCGAAGGATGGACATCCTGCAATGGATTGAGCCTG
CTGGTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTGAGCTTAGTTGAGGAAGACCAATTT
TCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATTATAGGAGCAGGTTCTGATGGCCATTCAGCA
ACAACATGTCTCTTGACAGCAAGAAAAAGAGCTGTGCAACAACAGAACTGGAATGTTTCTTTATCATCTTTTC
AAGTGATACACAGTCATTGGTGCTCTGTATTGCGATGCTGATATCCATCCAGGCTCTCTTTAAAAAGTGCTCTCAATG
TGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTTCATCCAGAATCC
TTCAACTTGCAGTGGTTTTTCAATGACTCTTGTCACCTCCTACTGGTTTCAATAAACCCACCAGTAACGACACC
ATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCA
GTATTTTTAGGTCTATTGCTTGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCCTT
GGCTGTCTAGTCTGGAGTCTCTAAGCGAAGAAGTCAAATTGTG**TAG**TTTAATGGGAATAAAATGTAAGTATCAGTA
GTTTGAAAAA

199/615

FIGURE 198

MTCCEGWTS CNGFSL LVL LLLGV LNAI PLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPAT TMSLTARKR
 ACCNNRTGMFLSSFFSVITVIGALY CMLISTIQALLKGPLMCNSPNSNANCEFS LKNISDIHPESFNLQWFFNDS
 CAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRR
 SQIV

200/615

FIGURE 199

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGAAGGAGCAGTG
AGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAGATCCGTGGGCTGCAGACCCC
CGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCT
ACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAA
AAACCTGCAGCTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAA
ATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTC
TGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACCTCCTC
CCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCAAGGGCAGGCTGTTAGGCCCCCTTTCTGATCAGGAGGC
TTCTTTATGAATTAACTCGCCCCACCACCCCTCA

200/615

201/615

FIGURE 200

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKSSQKQHSPVPE
KAIP LITPGSATT C

201/615

202/615

FIGURE 201

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGGAGGACAGGGA
GTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGGAGGAGACCCCTGGTGGGAGGA
AGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAGGGGCCCCCTGGCCTGCCTCCTGCTGGCCCTC
TGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAGCGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTT
GGACATGGCCTGGGAGACGCCCTGAGCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGC
TCTAAAGTCAGTGAGGCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCAGGCTTT
GGCGCAGCAGATGCTTTGGGCAACAGGGTCCGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGATTGGC
AGACAGGCAGAAGATGTCATTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGGGTGCCTGGCCACAGT
GGTGCTTGGGAAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGGCCTTGGAGGCCAGGGCCAGGGCAAT
CCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCGGAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAG
GGAGCTCCCTGGGGTCAAGGAGGCAATGGAGGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAG
CCTGGCTATGGTTTCAGTGAGAGCCAGCAACCAGAAATGAAGGGTGCACGAATCCCCCACCATCTGGCTCAGGTGGA
GGCTCCAGCAACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGGCAGC
AGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCCTCCTGGGGATCCAGCACC
GGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGACATAAACCCGGGTGTGAAAAGCCAGGGAAT
GAAGCCCGCGGAGCGGGGAATCTGGGATTCAGGGCTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATA
AGCAAAGAGGGCAATCGCCTCCTTGGAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGGA
GGAGGTGACGCTGTTGGTGGAGTCAATACTGTGAACTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTC
TGGAAGAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAAGGACCAGAGAAGCTCTCGCATC
CCGTGA CCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAAACACCACCCTCTCA
TCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA
AAAAAAAA

202/615 " 202/615"

203/615

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212
><subunit 1 of 1, 440 aa, 1 stop
><MW: 42208, pI: 6.36, NX(S/T): 1
MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVSEALGQGT
REAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVPGHSGAWETSGGHGI
FGSQGGLGGQGGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQ
NEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSGNSGGS
RGDSGSESSWGSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRRGQGVSSNMREISKEGNRLLGGS
GDNYRGQGSWSGSGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQRSSRIIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80, 90-96,
96-102, 130-136, 140-146, 149-155, 152-158, 155-161, 159-165, 163-169,
178-184, 190-196, 194-200, 199-205, 218-224, 236-242, 238-244, 239-245,
240-246, 245-251, 246-252, 249-252, 253-259, 256-262, 266-272, 270-276,
271-277, 275-281, 279-285, 283-289, 284-290, 287-293, 288-294, 291-297,
292-298, 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 203

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTGCTGCTGAGCCTGCCCTGGCTGGGCCTCA
GACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCTACTCGCCCGCATCCTGGCTT
GGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCCACAGCCCCCAAACGGAACTGGTTTTTGGG
GTCACCTGGGCCTGATCACTCCTACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGG
GCTTTACGGTATGGCTGGGTCCCATCATCCCTTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCA
ATGCCCTCAGCTGCCATTCACCCCAAGGATAATCTCTTCATCAGGTTCCCTGAAGCCCTGGCTGGGAGAAGGGATAC
TGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCTGAAGT
CCTATATAACGATCTTCAACAAGAGTGCAAAACATCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCA
GTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACA
GCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCC
AGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCC
TGGTGCATGACTTCACAGACGCTGTCTCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTT
TCAAAGACAAAGCCAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTGAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCC
TCTCCTGGGTCTGTACAACCTTGCAGAGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGCAAGAGCTTC
TGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCAGCTGCCCTTCCCTGACCATGTGCGTGAAGG
AGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCC
GAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATC
CTGAGGTCTACGACCCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCCTTTCT
CCGCAGGGCCAGGAAGTGCATCGGGCAGGCGTTCCGCATGGCGGAGATGAAAGTGGTCTGGCGTTGATGCTGC
TGCACTTCCGGTTCTTGCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGCGCGCCGAGGGCGGGC
TTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCAGTGACTTTTCTGACCCATCCACCTGTTTTTTTTTGCAGATT
GTCATGAATAAAACGGTGCTGTCAAA

205/615

FIGURE 204

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLGLITPTEGLK
DSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKPWLGEIGILLSSGDKWSRHRM
LTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHISLMTLDSLQKCIFSDSHCQERPSEYIATI
LELSALVEKRSQHILQHMDFLYYLSHDGRRFHRACRLVHDFTDVIRERRRTLPTQGIDFFKDKAKSKTLDLDFID
VLLLSKDEDGKALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDL
AQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENS
KGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLNVGLQ

205/615

FIGURE 205

TCCCTTGACAGGTCTGGTGGCTGGTTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAACTGAAGACTCTCTGCT
 TTTGCCACAGCAGTTCCCTGCAGCTTCCCTTGAGGTGTGAACCCACATCCCTGCCCCCAGGGCCACCTGCAGGACGC
 CGACACCTACCCCTCAGCAGACGCCGGAGAGAAATGAGTAGCAACAAAGAGCAGCGGTGAGCAGTGTTCGTGATC
 CTCTTTGCCCCTCATCACCATCCTCATCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTG
 CGGGGCCGTAGCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGCAAC
 AAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTGATGAGCTCCAGCCACCTGCTGGGCACCAAGCTGGGC
 CCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCACCCACCACTGGCTACTCAGCTGATGTGGGC
 AACAAGACCACCTACCGCGTCTGTGGCCCATTCAGTGTGTTCGCGTGTCTGAGGAGGCCCCAGGAGTTTGTCAAC
 CGGACCCCTGAAACCGTGTTCATCTTCTGGGGGCCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCCT
 GTGATCCAGCGAGCGGGCCTGGTGTTCCTCAACATGGAAGCATATGCCGTCTCTCCCGGCCGATGCGGCAATTT
 GACGACCTCTTCCGGGGTGTAGACGGGCAAGGACAGGGAGAAGTCTCATTCTGTTGGTGTGAGCACAGGCTGGTTTACC
 ATGGTGTATCGCGGTGGAGTTGTGTGACCACGTGCATGTCTATGGCATGGTCCCCCCCCAACTACTGCAGCCAGCGG
 CCCCCTCCAGCGCATGCCCTACCACTACTACGAGCCCAAGGGGCCGACGAATGTGTCACTACATCCAGAAT
 GAGCACAGTCGCAAGGGCAACCAACCCGCTTCATCACCGAGAAAAGGGTCTTCTCATCTGTTGGGCCAGCTGTAT
 GGCATCACCTTCTCCACCCCTCCTGGACCTAGGCCACCCAGCCTGTGGGACCTCAGGAGGGTGTGAGGAGAAGC
 AGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTGCTGGAGTGTCTCCAGCCAATC
 AGGGCCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCTTGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGT
 CTATGTGGTTAATCAGGGGTGTCTTCTTGTGTGAGTGTGAGGCTGTGCGCACAGTCAATCAGGGTAGAGGGGGTATT
 TCTGAGTCAATCTGAGGCTAAGGACATGTCTTTCCCATGAGGCCTTGGTTTCAAGGCCCCAGGAATGGACCCCC
 AATCACTCCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACCTGGTGTGCCCCCTCAATTT
 CCAGCACCAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCCCGGCCAGAGAATTTGTGGGGTTGTGG
 AGGTTGTGGGGGGCGGTGGGGAGGTCCCAGAGGTGGGAGGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTG
 GACAAACCCCTCCCCCTCTCTGGGCACCCCTTCTGCCCACACCAGTTTCCAGTGCGGAGTCTGAGACCCCTTCCAC
 CTCCCCCTACAAGTGCCCTCGGGTCTGTCTCCCCGTCTGGACCCCTCCCAGCCACTATCCCTTGCTGGAAGGCTCA
 GCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACCTTAGGGTATTTTTCGCAAACTCCTTCAGG
 GTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCTTAGCCCCCTCAGCCAGCTGCCATTAGCTT
 GGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCTCTAGCAGGGAGGTTTTTCCAAGTGTGAGGCGCCTTTGGGG
 CTGCCCTTTGTCTGGAGTCACTGGGGGCTTCCGAGGGTCTCCCTCGACCCTCTGTCTGCTGGGATGGCTGTCTG
 GGAGCTGTATCAGCTGGGTTCTGTCCCTGGCTCTGTATCAGGCACTTTATTAAGCTGGGCCTCAGTGGGGTGT
 GTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGAGGCTGGAGGGACCAGATG
 GAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGGGGCGGTGACTGCCCCAGACTTGGTTTTGTGA
 ATGATTTGTACAGGAATAAACACACCTACGCTCCGGAAAAA

207/615

FIGURE 206

MSSNKEQRSVAVFVILFALITILILYSSNSANEVVFHYGSLRGRSRRPVNLKKWSITDGYVPILGNKTLPSRCHQCV
IVSSSSHLLGTKLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAHSSVFRVLRRPQEFVNRTPETVFI FWG
PPSKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRMRQFDDLFRGETGKDREKSHSWLSTGWFTMVI AVELCDHV
HVGVMVPPNYCSQRPLQRPYHYEYPKGPDECVTYIQNEHSRKGNHHRFITEKRVFSSWAQLYGITFSHPSWT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 9-31 (type II)

N-glycosylation site.

amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 50-54

Casein kinase II phosphorylation site.

amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.

amino acids 253-262

N-myristoylation site.

amino acids 37-43, 114-120, 290-294

FIGURE 207

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCTGGAGCCAGGAGCGACGTCAC
CGCC**ATGG**CAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTTGG
ATGTGCCCTTCCAATATACAACAAATACTGGCCCTCTTTGTTCTATTTTTTACATCCTTTCACCTATTCCATA
CTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGCCATCTTCTTAC
AACGGGCATTGTCTGTCTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGAGCTTG
TGCACTTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAGCAATGA
CGACTTCAGCTGGCAGCAGTGG**TGA**AAAAGAAATTACTGAACCTATTGTCAAATGGACTTCCTGTCAATTTGTTGGCC
ATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTAGGTGCTCC
CTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAGGATTTTCTCTTTTGGAA
AAGCTTGACTGATTTTCACTTATCTATAGTATGCTTTTTGTGGTGTCTGCTGAATTTAAATATTTATGTGTTT
TTCCTGTTAGGTGATTTTTTTTTTGGAAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCATTTGCATGGT
TAGGAATTCAGAAATCCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAATTTATTTAGCCTCCATTA
TTACAAAAAATTATAAAAAATAAGTTTTTCAGTCAGTCAGGATGACATCACTCCAATGTTATGCAGACATACAGAC
GGTTGGCATACTGTTATAGACTGTATACTCAGTGCAAAATATAGCTGCATTTATACCTCAGAGGGGCCAAGTGTTAA
TGCCCATGCCCTCCGTTAAGGGTGTGGTTTTACTGGTAGACAGATGTTTTGTGGATTGAAAATTTATTTATGG
AATTGCTACAGAGGAGTGCTTTTCTCTCAATTGTTAGAAGAATTTATGTTAAACTTTAAGGTAAGGGTGTAAAA
ACATTTTTGAGATAAGGTTTTTATTTATGTTTATTTATTTAGAGTGAGTTGCAATGTGGGAAGAAATGACATTG
AAATTCAGTTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATCAACCTTTTCATGTTTTACC
CTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGCATCATATATGCCAGAAAAAC
CTTCTCTGCTTCTCTCTTTTGAAGTTATTTGGTATGTTGTATATATTACATAAAAATAAAGTTTTCAAATATAGTTT
AATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATTCAGAGTGCCCCCTCCCCTGCA
AGGCCTTGCCATGATTAAACAAGTAAGTTGTTAGTCTTACAGATAATTTCATGCATTAACAGTTTAAAGATTTAGACC
ATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAGTATTTTAAAGACAAGTTTCCTGT
ATACCTCTGAAGTGTGTTGATTTGAGTTTCATCATGATAGATCTGCTGTTTCTTATAAAAGGCATTTGTTGTGT
GAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACATACCTGACCAAAAAATTTCCAGTAAC
CAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGACTTTTTTCAGGAGTGGGTTATAAAAA
CATTCAGTTGGTCTGACAGTATTTTTGTTAAGGATATTTGTTTGTATGTTTATTCAGTATACTTACATAAAAAAT
ATTTGCCATCAGCCAAAAGTCAAGTAAATCATGACAGCTGCTGTTGTTTATGAAAGTTTATTTCTCAAGAAAATG
GGAATAAATTTGGGATTTGTTTCAAGCTTTTTTACTAAAGATGCCTAAAGCCACAGGTTTTATTGCCTAACTTAAGC
CATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCGGCGTGTGGCTGGAGCCTTCCCCTGGAGGC
TGAAAGTGGCTTGTGGTATTATAATGTTTCAGATTTCAAGAGGAAGGTGCAGGTACACATGAGTTAGAGAGCTGGT
GAGACAGTTGGGAAGTCTTTGTGCTTGTGATCTACTGGACTTTTTTTTTTGCAGGAAGTGCATTTCTCTGGTCCTTC
CCTATTTTCTGTTCTGGATGTCAGTGCAGTGCAGTGCCTACTGTTTTATCCACTTGGCCACAGACTTTTTCTAACA
GCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATTGTGTCTTTGACCTTGTATACTAGCTTGACATAGT
GCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAATTTTCCATAGAATATGCACTGATACAAACATTACCAT
TCTTCTATGGAAAGAAAAGCTTTTGATGATGAAACAATAAAGATTTTAAATATCTATTTTAAAAA

209/615

FIGURE 208

MAGIKALISLSFGGAIGLMFLMLGICALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAMSNACKELAIFLT
T
GIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSNDDFSWQQW

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

210/615

FIGURE 209

CTTGCAGAGAAAGAGTCTTTTGTGCAGCACCCCTTTAAAGGGTGACTCGTCCCCTTGTGTTCTCTCTCCTGGTGC
AGAGTTGCAAGCAAGTTTATCAGAGTATCGCCATGAAGTTTCGTCCCCTGCCTCCTGCTGGTGACCTTGTCTCTGCC
TGGGGACTTTGGGTCAGGCCCCGAGGCAAAAGCAAGGAAGCACTGGGGAGGAATTCCATTTCCAGACTGGAGGGA
GAGATTCTTGCACTATGCGTCCCAGCAGCTTGGGGCAAGGTGCTGGAGAAGTCTGGCTTCGCGTCGACTGCCGCA
ACACAGACCAGACCTACTGGTGTGAGTACAGGGGGCAGCCAGCATGTGCCAGGCTTTTGCTGCTGACCCCAAAC
CTTACTGGAATCAAGCCCTGCAGGAGCTGAGGCGCCTTCACCATGCGTGCCAGGGGGCCCCGGTGCTTAGGCCAT
CCGTGTGCAGGGAGGCTGGACCCCAGGCCCATATGCAGCAGGTGACTTCCAGCCTCAAGGGCAGCCAGAGCCCA
ACCAGCAGCCTGAGGCTGGGACGCCATCTCTGAGGCCCAAGGCCACAGTGAAACTCACAGAAGCAACACAGCTGG
GAAAGGACTCGATGGAAGAGCTGGGAAAAGCCAAACCCACCACCCGACCCACAGCCAAACCTACCCAGCCTGGAC
CCAGGCCCGGAGGGAATGAGGAAGCAAAGAAGGCCCTGGGAACATTGTTGGAAACCTTCCAGGCCCTGTGCG
CCTTTCTCATCAGCTTCTTCCGAGGGTGAAGGTTGAAAGACCCCTACAGATCTGACCTCTCCCTGACAGACAACC
ATCTCTTTTTATATTATGCCGCTTTCAATCCAACGTTCTCACACTGGAAGAAGAGAGTTTCTAATCAGATGCAAC
GGCCCAAATTCCTTGATCTGCAGCTTCTCTGAAGTTTGGAAGAAACCTTCTTTCTGGAGTTTGCAGAGTTCAG
CAATATGATAGGGAACAGGTGCTGATGGGCCCAAGAGTGACAAGCATACACAACCTACTTATTATCTGTAGAAGTT
TTGCTTTGTTGATCTGAGCCTTCTATGAAAGTTTAAATATGTAACGCATTTCATGAATTTCCAGTGTTTCAGTAAAT
AGCAGCTATGTGTGTGCAAAATAAAAGAATGATTTTCAGAAAAAAA

210/615

211/615

FIGURE 210

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59602
<subunit 1 of 1, 223 aa, 1 stop
<MW: 24581, pI: 9.28, NX(S/T): 0
MKFVPCLLLVTLSCGLTLGQAPRQKQGSGTEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR
VDCRNTDQTYWCEYRGQPSMCQAFADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA
GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRFPKATVKLTEATQLGKDSMEELGKAKPTTR
PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG

Important features:

Signal peptide:

Amino acids: 1-19

N-myristoylation sites:

Amino acids: 38-44;51-57;194-200

DNA photolyases class 1 proteins:

Amino acids: 58-69

Tyrosine kinase phosphorylation site:

Amino acids: 64-71

N-myristoylation sites:

Amino acids: 38-44;51-57;194-200

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids: 4-15

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

212/615

FIGURE 211

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCGGGTCCTGCTGTGGCTGCAGCTCTGCGCACTGACCCAGGCG
GTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAACTGGAGCCAGAACCGGACCCCGTGCGCC
GGCGGCGCCGTTGAGTTCCCGGCGGACAAGATGGTGTGAGTCCCTGGTGCAAGAAGGTACAGCCGTCTCAGACATG
CTCCTGCCGCTGGATGGGGAACCTCGTCCTGGCTTCAGGAGCCGGATTCCGGCGTCTCAGACGTGGGCTCGCACCTG
GACTGTGGCGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCGCACCTGTGGCGCT
CTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGCCACGACGACGTCTTCTTTC
CGCCTAGTGCCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAGCCCCGTGCGTGTCCGCAGCATCTCGGCTCTGG
GCCGGACGTTACGCGCGACGAGGACCTGGCTGTTTTCTGGCGTCCCGCGCGGGCCGCTACGCTTCCACGGGC
CGGGCGCGCTGAGCGGTGGGCCCCGAGGACTGCGCGGACCCGTGCGGCTGCGTCTGCGGCAACGCGGAGGCGCAGC
CGTGGATCTGCGCGGCCCTGCTCCAGCCCCCT

CGTGGATCTGCGCGGCCCTGCTCCAGCCCCCT

213/615

FIGURE 212

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59603
<subunit 1 of 1, 197 aa, 1 stop
<MW: 20832, pI: 8.74, NX(S/T): 2
MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLV
QEGHAVSDMLLPDGLVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDRTCGAL
GTRHLASSSWTPSACPAATTTSSFRLVPPSAWGSALALAPCVSAASRLWAGRSRATRTWL
FSWRPARAAYASTGRAR

Important features:

Signal peptide:

Amino acids 1-19

N-glycosylation site:

Amino acids 35-39

Glycosaminoglycan attachment site:

Amino acids 81-85

N-myristoylation sites:

Amino acids 82-88;118-124;153-159

C-type lectin domain proteins:

Amino acids 108-118

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 213

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGAGCCTACTCGT
TTGATTGCAACTATCATGGTGCCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCTTTTGGTGGCATAACAAGGGA
CTTGCACTTATCTTCTGCATTTTGCAGTCCTTTGGCATTGACGTGGTACAGCCTTTCCTTCATACCATTTGCAAGG
GATGCTGTGAAGAAGTGTTTTGCCGTGTGTCTTGCATAATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTA
TGGACAGAAGCTGGTGGACAGTTTTGTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGC
AGCAATGTGTTGCTTGTGATTGGAACATTTGAGGGTTACTTTTGGGAAGCAACAATACATTCTCGAACCTGAATGT
CAGTAGCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAATCTTCCTCATGTACCTGTTTCCTCTCTGG
ATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

215/615

FIGURE 214

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFARDAVKKCFVCLA

2025.04.10 09:59:00

[illegible]

GGATTTTTTGTGATCCGCGGATTTCGCTCCCAACGGGCGGGACCTTTGTAACTGCGGGAGGGCCCAGGACAGGCCACCC
TGCGGGGCGGGAGGCAGCCGGGTGAGGGAGGTGAAGAAACCAAGACGCAGAGAGGCCAAGCCCCCTTGCCCTTGGG
TCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAACCAGATCCAGAGGCAACAGGGGACATGGCCACCTGGGACG
AAAAGGCAGTCAACCCGACGGGCCAAGGTGGCTCCCGCTGAGAGGATGAGCAAGTCTTAAAGGCATCTCACGGTGC
TGGGAGACGACTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAAATGAAGAGGAGGAGGAGGAGGAGC
AGCCACCACCCACACCAAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCGTGACGTTGCCCTGCCCTGGCCCCG
CAGCCAGGGCCCCCTTGACTTCAGGGGCATGTTGAGGAAACTGTTGAGCTCCCAAGGTTTCAGGTTCATCATCA
TCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGCTTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACA
AGAATAACATGATGCTGCCATGGTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCT
TTAAATATTATTGTCTTCCGCTGTAGTCTTTTACCACAAGTTTGAGATCTTGATGCCCCGTGTTGGTGGTGGTCT
CATTCATCCTGGACATTGTCTCTGTTTCCAGGAGCACCAAGTTTGAGGCTCTGGGCCGTGCTGATTCGTCTGCTCCGCG
TGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGTTAAGACACGTTCAGAACCGCAACTCTTAAAGGT
TAAACAGATGAATGTACAATTGGCCGCCAAGATTCAACACCTTGAGTTGAGCTGCTCTGAGAAGCCCCCTGGACT
GATGAGTTTGTCTGTATCAACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTA
CTCTCACACAGCCACCGTGAAAGTCTTGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAACAGGCTGGC
ATGTTCACTGGGCTGGTGTTTACGACAGAGAACCTGCAGCTCACTGGCCAGTTATCATCTCAGATTACAAATCACA
CAGAGCATCTGCTGTTTCAATCAAGAGAAACAAAACCTATAAAGATATTCTGAAAATATGACAGAA
TTTGACAAATAAAAGCATAAACGCTGTAAAAA

217/615

FIGURE 216

MATWDEKAVTRRAKVAPAERMSKFLRHFTTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPVSGEEGRAAAPDVA
PAPGPAPRAPLDFRGMLRKLFSSSHRFQVIIICLVVLDALLVLAELILDKIIQPDKNNYAAMVFHYMSITILVFF
MMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILDIVLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSE
RQLRLKQMNVQLAAKIQHLEFSCSEKPLD

20570" 99525007

FIGURE 217

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGCTTCCCTGGGG
CAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTGCACTCATCATTTGGCTTTGGT
ATTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGC
TGCACTTTTGAACCTGACATCAAACCTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTC
CATGAGTTCAAAGAAGGCAAAGATGAGCTGTGCGAGCAGGATGAAATGTTTCAAGGCCGGACAGCAGTGTGCT
GATCAAGTGATAGTTGGCAATGCCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGT
TATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACCTGGAGCCTTCAGCATGCCGGAAGTG
AATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTC
TGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAAT
GTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAAAATGAC
ATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAAC
TCAAAGGCTTCTCTGTGTGTCTCTTCTTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATG
CTAAAATAATGTGCCTTGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAG
AGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAATCTATCTTCAAAGACATATTA
GAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAGTAAAATGCACGTGGAGACAAGT
GCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTG
TCTCTGAATTTTTTAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCA
CATCTTATATTCCACAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAG
GGGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTCTCTTC
CCAAC TGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGTCGGGGACACCCGATTT
TATAAATAAACTGAGCACCTTCTTTTTTAAACAAA
AAAAAAA

219/615

FIGURE 218

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLSDIVIQWLKEG
VLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKCYIITSKGKGNANLEYKTGAF
SMPEVNVDYNASSETLRCEAPRWFPPQPTVVWASQVDQGANFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSC
MIENDIAKATGDIKVTESIKRRLSHLQLLSKASLCVSSFFAISWALLPLSPYMLK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted November 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 219

GAATTTGTAGAAGACAGCGGCGTTGCC**AT**GGCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGTGCTGGTGGCCGCT
CTGTGGGGTGGCAGCGAGCCGCTGCTGAAGCGGGCCCTCCGCCGGCCCTGCAGCGGGTTCATGAGCCGACCTGGGCC
CAGCAGTTGCTACAGGAGATGAAGACCCCTCTTCTTGAATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGT
GGATCCCTTCTCTATTACCTCACCTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCT
ATCATCTTCACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGCGAG
TGCGGGACGCAGCTCTGTGGATCTCGACATACCTGTGTTAGTTCCCTTCCAGAACCCATCTCCCCAGAGTGGGTG
AGGACACGGCCCTTTTCCCATCCTGCCCTTTCCCTCTGCAGCTGTTTTGCTTCCCTTGTGGCCATCAGAGTTCCCTTC
CCCTGGACAGTCTGGAGAAAGACAGAGGCTGGGGTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCCTCCA
GCCCTGTACCAGCTCCTACTGGCATGGCTGAGCTCAGACCCCTCCTGATTTCTGCCATTATTATCCCAGGAGCAGTTG
CTGGCATGGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGCAACAGT
CTACCCCTTTGAGTGGGCGCAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCCTGGGCCATGAAGTGCTGGCA
GTGAGCGGATGGACCTAGCACTTCCCCCTCTCTGGCCTTAGCTTCCCTCCTCTCTTATGGGGATAACAGCTACCTCA
TGGATCACAATAAGAGAACAAGAGTGAAAGAGTTTGTAACTTCAAGTGCTGTTTCAGCTGCGGGGATTTAGCAC
AGGAGACTCTACGCTCACCCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCTGCTAACATCTCAGGCTCCAG
CCCAGCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACTCCAGCTGCA
TGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTGGGACTCAGTCAGAGCGCTTTGG
CTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGAGTGGCAGATGCAGGAAATGAGCTGTCTATAGCCT
TGCCTGCCCCACCCATGAGGTAGGCAGAAATCCTCACTGCCAGCCCCCTTTAAACAGGTAGAGAGCTGTGAGCCC
CAGCCCCACCTGACTCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAA
AA

221/615

FIGURE 220

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMFPFLNQCGSLLYYLTL
ASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISP EWVRTRPFPILP
FPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

222/615

FIGURE 221

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTGGAAGATGACAGCA
ATTATAGCAGGACCCTGCCAGGCTGTCGAAAAGATTCCGCAATAAACTTTGCCAGTGGGAAGTACCTAGTGAAA
CGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCTTGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTC
CAGCACCATGAAAGGGCATCCTCGTTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGT
GCAGTGTAATTCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAGCTG
TATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGTTCTGCTCAGCGGAGAA
CTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCTGCTGAAGAACAATTTTCAATTTTGTAAAGCCA
GTGCTGCCAAGGAAAGGAATGCAGCAACACCAGCGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGC
AGAGTGCCCTGCTTGTTATGAATCTAATGGAACCTCCTGTCTGGGAAGCCCTGGAAATGCTATGAAGAAGAACA
GTGTGTCTTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTCCAACGT
CAGTAACGCCACCTGTCTAGTTTCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCATCTTTGAAAAGTTTGAGTG
TGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACTTCCCACAACGTGGGCTCCAAAGCTTCCCTCTA
CCTCTTGGCCCTTGCCAGCCTCCTTCTTCGGGGACTGCTGCCCTTGAGGTCCCTGGGGCTGCACTTTGCCAGCACC
CCATTTCTGCTTCTCTGAGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTCCCTGCTCTGCCCCGTTTAA
CTGCCCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTAAAGCACTGG
TTCATTCACTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

205110 "98525001

223/615

FIGURE 222

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVRLYQNMFCSAENC
EETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAECPACYESNGTSCRGKPWKCYEEEQCV
FLVAELKNDIESKSLVLKGCSNVS NATCQFLSGENKTLGGVIFRKFE CANVNSLTPTSAPTTS HNVGSKASLYLL
ALASLLLRGLLP

2025-05-20 10:00:00

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCAGGGCTTGCCCTCAC
TGGCCACCCCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCGGCGCGCTGCTGTGGGTCCCTGCTGCTG
AATCTTGGGTCCCCGGGGCGGCGGGGGCCCAAGGCCCTGACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGC
TTTGGGGGCCCTATGACCCCGAGCTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTTA
GAGGACGAGAATGATGCCATGGCCGACCGCGACCGCTGGCTGGACAGCGGGTCCCGAGCTCTTGGCCGCCACG
GTGTCCACCGGCTTTAGCCGGTCTGCCGCCATTAAAGAGGAGGATGGGTCTTCAGAAGAGGGGGTTGTGATTAAT
GCCGAAAGGATAGCACCAGCAGAGAGCTTCCAGTGCGACTCCCAATACAGCGGGGAGTTCCAGCACGAGGTTT
ATAGCCAATAGTCAGGAGCCTGAAATCAGGCTGACTTCAAGCCTGCCGCGCTCCCCGGGAGGTTCTACTGAGGAC
CTGCCAGGCTCGCAGGCCACCCCTGAGCCAGTGCTTCCACCTGGGTCTACCCCGAGCCGGTGGCCGTCAACCTCA
CCACAGCCATGCCATCTCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCACTGCAAG
TCGGGCACCATGAGCCGGAGCCGTTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCTTCGAGTTGGGGCGCTGAGC
CAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTCCCTGCAACCGACTTCGGAAGAGTGCCCCCTG
GACACAAGTCTCTGTACTGACACCAACTGTGCCTCTCAGAGCACCACAGTACCAGGACCACCACTACCCCTTC
CCACCATTCACCTCAGAAGCAGTCCGAGCCTGCCACCCCGCCAGCCCTTGCCAGCCCTGGCTTTTTTGGAACGG
GTCAGGATGGCCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTTACAGAGATGCAACCAATAGACAGAAAC
CAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAATCTCAACCTCTCTTGCCCTTCAATCCTAGCAC
CCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAAACAA

[illegible]

MVPAAGALLWVLLLLNLGPAAAGAQGLTQTPTTEMQRVSLRFGGPMTRS.YRSTARTGLPRKTRIILEDENDAMADAD
 RLAGPAAAEELLAATVSTGFSRSSAINEEGDSSEEGVVINAGKDSTSRELPSATPNTAGSSSTRFIANSQEPETRL
 TSLSPRSGRSTEDLPGSQATLSQWSTPGSTPSRWSPSPSTAMPSPEDLRLVLMPPWGPWHCHCKSGTMSRSRSGK
 LHGLSGRLRVGALSQRLREHKPCTYQCPNRLREECPLDTSLCDTNCASQSTTSTRTTTTTFPPTIHLRSSPSL
 PPASPCPALAFWKVRVIGLEDIWNLSLVFTEMQPIDRNQR

FIGURE 225

CCCGGGTCGACCCACGCGTCCGGGGAGAAAGG**ATG**CCCGCCTGGCGGCGCGGTTGGTCCTGCTAGCTGGGGCAG
 CGGCGCTGGCGAGCGGCTCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAAGAGCAGA
 ACTGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCCGCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCTGTC
 GGGACGACTGTAAGTATGAGTGTATGTGGGTACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCTCAGT
 TCCATGGCAAGTGGCCCTTCTCCCGGTTCTGTTCTTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCTCAATG
 GCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCATGTACCACACCTGTG
 TGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGACCTCACAG
 AGAAATGGACTACTTCTGTGCCTCCACTGTCACTACCTCAATCTACCTGTGCTGCGTCAGGACCGTGGGGC
 TGCAGCACCCAGCTGTGGTCACTGTCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCTCCTACCTGA
 GCCTCATCCGCTTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTGGTGTGGTGGC
 TGGCCTGGTGCCTGTGGAACAGCGGCGGCTGCCTCACGTGCGCAAGTGCCTGGTGGTGGTCTTGCTGCTGCAGG
 GGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCTGGATGCCCATGCCATCTGGCACATCA
 GCACCATCCCTGTCCACGTCTCTTTTTCAGCTTTCTGGAAAGATGACAGCCTGTACCTGCTGAAGGAATCAGAGG
 ACAAGTTCAAGCTGGACT**TGA**AGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCCGCCCTGCTGGCCTC
 CCTTCTCCCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAAGTGGACATGAAGGATGTGGGGCCAG
 AATCATGTGGCCAGCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGGAAGGCCTCCAGCATC
 TGGGACTCGAGAGTGGGCAGCCCTCTACCTCCTGGAGCTGAAGTGGGGTGGAACTGAGTGTGTTCTTAGCTCTA
 CCGGGAGGACAGCTGCCGTGTTTCTCCCCACAGCCTCCTCCCCACATCCCCAGCTGCCCTGGCTGGGTCTTGAAG
 CCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCCTTCTGTTACCAACCCCCCA
 CCCTCCTCCAGGACACCCTAGGTGGTGTGGATGCTTGTCTTTTGGCCAGCCAAGGTTACAGGCGATTCTCCCC
 ATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTACCCTGACCGTTGCCCTAGCCAGGTTCCCCA
 GGAGGCCTCACCATACTCCCTTTTCCAGGGCCAGGGCTCCAGCAAGCCAGGGCAAGGATCCTGTGCTGCTGTCTGG
 TTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCCAGGCTGAGTGCATAGGTGACAGGGCCGTGAGCATGGGCC
 TGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGTGCGGGGAAGAGGTGTGGCTTCAAAG
 TGTGTGTGTGCAGGGGGTGGGTGTGTTAGCGTGCGTTAGGGGAACGTGTGTGCGCGTGTGGTGGGCATGTGAGA
 TGAGTGAAGTCCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGATGAGGGAATCCTGTACCATCAATAAT
 CACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGCCAGGAGCTCTCCATGGCCAGGCTGCCT
 GTGTGCATGTTCCCTGTCTGGTGCCCCCTTGGCCGCTCCTGCAAACCTCACAGGGTCCCCACACAACAGTGCCC
 TCCAGAAGCAGCCCCCTCGGAGGCAGAGGAAGGAAAAATGGGGATGGCTGGGGCTCTCTCCATCCTCCTTTTCTCCT
 TGCCTTCGCATGGCTGGCCTTCCCCTCCAAACCTCCATTCCCCTGCTGCCAGCCCCCTTGGCATAGCCCTGATTT
 TGGGGAGGAGGAAGGGGCGATTTGAGGGAGAAGGGGAGAAAGCTTATGGCTGGGTCTGGTTTCTTCCCTTCCCAG
 AGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGAGGGGCCACACTATGCCTGTGCCCTGGTAAAGGTGACCC
 CTGCCATTTACCAGCAGCCCTGGCATGTTCTGCCCCACAGGAATAGAATGGAGGGAGCTCCAGAACTTTCCAT
 CCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTGTCTGCCCCTGACCCCTTGTCCCTCTTTGAGGGA
 GGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTGGCCTGCGCTAGCTTCTTTTGATACTGAAAAC'TTT
 AAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAATTCCAAGCCTCAAAAAAAAAAAAAAAAAA

227/615

FIGURE 226

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMFLAGWTCRDDCKYECMWV
TVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFLNGLASLVMLCRYRTFVPASSPMYHTCVAFWVSLNAW
FWSTVFHTRD'TDLTEKMDYFCASTVILHSIYLCCVRTVGLQHPAVVSAPFRALLLLMLTVHVSYLSLIRFDYGYNL
VANVAIGLVNVVWWLAWCLWNQRRLPVVRKCVVVVLLQLGLSLLLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFS
FLEDDSLYLLKESEDKFKLD

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

20250409 10:50:00

FIGURE 227

TTCGGCTTCCGTTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCCCTTCCCCGGGGTC
 TGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTTCGCGTTGCCACCCACGCGGACTCCCCAGCTGGCGCGCCCC
 TCCCATTGCTGTCTGGTCAGGCCCCCACCCTTCCCACCTGACCAGCC**ATG**GGGGCTGCGGTGTTTTTCG
 GCTGCACTTTTCGTCGCGTTTCGGCCCGGCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTA
 TCATCCTGGTCGAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCC
 ATGTGACCGACCGGTGAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTGTCTCTTCTAC
 AGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAAGGGTTAGCATCGCTGAGTGAGGACG
 GAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCT
 CTGTTATCAATATTTTGGCTGATGCACCTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCCATTACTTCC
 TGACTTCAGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATGCCTGTG
 AGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCGGGACTGACATTCCTGAACC
 CCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGTTTCCATGGGGCTCTGGGCCTTCATCACAGCTG
 GAGGGTCCCTCCGAAGTATTCAGCGCAGCCTCTTGTAAGGACT**TGA**CTACCTGGACTGATCGCCTGACAGATCC
 CACCTGCCTGTCCACTGCCCATGACTGAGCCCAGCCCCAGCCCGGTCCATTGCCACATTCTCTGTCTCCTTCT
 CGTCGGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTTTACCAGGAG
 CAGCCTGGGTTTCAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCCCACCACCTGGGGACCCCTTGT
 TGTGTCCAGGACTCCCCCTGTGTGCTGCTCTGCTCTCACCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGG
 CCGACGGCAGGAGGACAGTCGGGTGATGGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGG
 GGGGACCCCTGGGCCTGGGGTGGCCTCCTGATGTCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCA
 GGTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAAGGATAGATGAGC
 TCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTCAGGCCTGAGGGGGAACCATTTTTGG
 TGTGATAAATACCCTAAACTGCCTTTTTTTCTTTTTTGAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTTCT
 AACCTCCTTGGGCTATATTTCTCTCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCCGTCCCTTTCTCCTTGG
 TCCCAGACCTTGGGGGAAAGGAAGGAAGTGCATGTTTGGGAACGGCATTACTGGAACATAATGGTTTTAACCTCC
 TTAACCACAGCATCCCTCCTCTCCCCAAGGTGAAGTGAGGGTGCTGTGGTGAGCTGGCCACTCCAGAGCTGCA
 GTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGGGAGATTTTTTTGTAGTTTAAATGGGG
 TGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGTATCATTTTCTGCTGAGGGTGAGGTGTCCCATCCTTTAATC
 AAGGTGATTGTGATTTTGAATAAAAAAGAAATTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

229/615

FIGURE 228

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVVWFILVHVTDRSDARLQYGLLI FG
AAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLSFGIISGVFSVINILADALGPGVVG I H
GDSPYYFLTSAFLTAAILLHTFWGVVFFDACERRRYWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMG
LWAFITAGGSLRSIQRSLCKD

229/615

230/615

FIGURE 229

CGGGAGGCTGGGTGCTCATGATCCGGACCCCATTTGTGGCCTCTGCCCATCGCCTGCTCCTCCCAGGCTCCCGCG
GCCGACCCCGCGCAACATGCAGCCACGGGCCGCGAGGGTTCCCGCGCGCTCAGCCGGCGGTATCTGCGGCGTC
TGCTGCTCCTGCTACTGCTGCTGCTGCGGCAGCCCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCAGAG
CCCTCTCCACGCTGGGCTCCCCAGCCTCTTACCAACGCCGGGTGTCCCCAGCGCCCTCACTACCCAGGCCCTCA
CTACGCCAGGCACCCCCAAAACCTGGACCTTCGGGGTTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCGTGG
ACGGCCACAATGACCTGCCCCAGGTCTTGAGACAGCGTTACAAGAATGTGCTTCAGGATGTTAACCTGCGAAATT
TCAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGTGGGTGCCCAGTTCTGGTCAGCCTCCGTCT
CATGCCAGTCCCAGGACCAGACTGCCGTGCGCCTCGCCCTGGAGCAGATTGACCTCATTACCCGCATGTGTGCCT
CCTACTCTGAACCTCGAGCTTGTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCCTATTGGCG
TGNAGGGTGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCTACCTGA
CACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTTCAGACACCACATGTACACCAACGTCA
GCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAACCGCCTGGGCATGATGATAGATTTGTCCTATG
CATCGGACACCTTGATAAGAAGGTCCTGGAAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAG
CTGTGTGTGACAATTTGTTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGA
CACTGTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTGACCACATCA
GGGCAGTCATTGGATCTGAGTTTCATCGGGATTGGTGAAATTATGACGGGACTGGCCGGTTCCCTCAGGGGCTGG
AGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCTGAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTG
TCCTTCGTGGAACCTGCTGCGGGTCTTCAGACAAGTGGAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCG
TGGAGGCTGAGTTTCCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCTCAGAATGGACACCAGG
CTACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCTCAAATGCCCTCCCCATACCTTG
TTCCAGGCCTTGTGGCTGCTGCCACCATCCCAACCTTACCCAGTGGCTCTGCTGAACAGTCCGTCCCCGCAGA
GGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCTAGTTTATTACAAAGCATATGCTGAGAATAAACATGTTA
CACATGGAAAA

CGGGAGGCTGGGTGCTCATGATCCGGACCCCATTTGTGGCCTCTGCCCATCGCCTGCTCCTCCCAGGCTCCCGCG

231/615

FIGURE 230

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817
><subunit 1 of 1, 487 aa, 1 stop, 2 unknown
><MW: 53569.32, pI: 7.68, NX(S/T): 5
MQPTGREGSRALSRRYLRRLLLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPSALTTPGLTTPGTP
KTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQTSLDRLRDGLVGAQFWSASVSCQSQD
QTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSSQKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTC
STPWAESSTKFRHHMYTNVSGLTSGFEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNL
LNVPPDILQLLKNGGIVMVTLSMGVLQCNNLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDVSTY
PVLIEELLSRXWSEEEELQGVLRGNLLRVFRQVEKVVREESRAQSPVEAEFPYGQLSTSCHSHLVPQNGHQATHLEV
TKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:

Signal peptide:

amino acids 1-36

Transmembrane domain:

amino acids 313-331

N-glycosylation sites.

amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359,
357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 136-146

20250307 10:00:00

232/615

FIGURE 231

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGATTGGCAAGCGCTGG
CCACCTCCCCACACCCCTTGCGAACGCTCCCTTAGTGGAGAAAAGGAGTAGCTATTAGCCAATTTCGGCAGGGGCC
GCTTTTTAGAAAGCTTGATTTCCCTTTGAAGATGAAAGACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGG
AACTCGGGGCGATTGGCTGGGAACGTATCCACCCAAATGTCACCGATTCTTCCCTATGCAGGAAATGAGCAGAC
CCATCAATAAGAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAGAGG
GTTGCTCAACGCCCCGCCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTGGCGGAGGCGGGCGGAT
GATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCTTTCCCCGCCCCCTGAGACCCTGCAGCACCA
TCTGTCTATGGCGGCTGGGCTGTTTGGTTTGGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCG
GCCGCCCCGCTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCC
CCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTGTATGAGAAGAACCCAGACTCC
CATGGTTATGACAAGGACCCCGTTTGGACGTCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATC
CTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCCGGAAGCTGAG
AGGCTTGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAACCTGCTTCGACCCCGCAAGATCCAG
CTGCCAGAGGATGAGTGAACAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCTGCCTGCCATTCT
TGACCTCTTCTCAGAGCACCTAATTAAAGGGCTGAAAGTCTGAA

232/615

233/615

FIGURE 232

MAAGLFGLSARRLLAAAATRG LPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPEDENLYEKNPDSHG
YDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAERLVKYREANGLPIMESNCFDPSKIQLPEDE

20570" 95625001

FIGURE 233

GCGGCGGCTATGCCGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCCGGCGGCTGGTGCCTTGCAAGACCCCCA
 CGCGACAGCCTGCGGGAGGAACTTGTTCATCACCCCGCTGCCCTTCCGGGGACGTAGCCGCCACATTCAGTTCCGC
 ACGCGCTGGGATTTCGAGCTTCAGCGGGAAGGAGTGTCCATTACAGGCTCTTCCCAAAGCCCTGGGGCAGCTG
 ATCTCCAAGTATTCTCTACGGGAGCTGCACCTGTTCATTACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCA
 CCCTTCCTGCAGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAATCT
 TGAAGGAGCTCAGTAATGTCTCTCAGGGATCTTCTGCGCTCTCTCAACTTCATCGACTCCACCAACACAGTC
 ACTCCCACTGCCCTCCTTCAAACCCCTGGGTCTGGCCAAATGACACTGACCCTACTTTCTGCGCTATGCTGTGCTG
 CCGCGGGAGGTGGTCTGCACCGAAAACTCACCCCTGGAAGAAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTC
 TCTGTGCTGCTGAAGGCAGATCGCTTGTTCACACCAGCTACCACTCCAGGCAGTGCATATCCGCCCTGTTTGC
 AGAAATGCACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCCTGTCACTTGTATTTGATGCCCTTCATCACG
 GGGCAGGGAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCCTCACGGAGCCCTGCCCCCTGGCTTCA
 GAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACAACGAGACATTAGAGGTGCACCCACCCCGACC
 ACTACATATCAGGACGTCTATCCTAGGCACTCGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATC
 AACAACTCTCGAAACCTCAACATCCAGCTCAAGTGAAGAGACCCCCAGAGAATGAGGCCCCCCCAGTGCCCTTC
 CTGCATGCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTACAACACCCAC
 CCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCGGCTGTATGTGCACACCCCTCACC
 ATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCCACTACCAGCCTGCCAGGACCGGCTGCAACCCAC
 CTCTGGAGATGCTGATTCACTGCTGCCGGCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTG
 AAGTGGACCGAGTACACGCCAGATCCTAACCATGGCTTCTATGTGAGCCCATCTGTCTCAGCGCCCTTGTGCCC
 AGCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCTCTTCAACAGCCTGTTCCCACTCTCTGATGGC
 TCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCCGACACCGGACTTCAGCATGCCCTAC
 AACGTGATCTGCCCTCAGTGCAC'TGTGGTGGCCGTGTGCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTC
 CACATCGAGGAGCCCCGCACAGGTGGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCC
 CCACTCTGATTCTTGCCCTTTCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTTTC
 TGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAGTGCCCTGGACCAGGTCAGGGCCTACAGCTGTGTTGT
 CCAGTACAGGAGCCACGAGCCAAATGTGGCATTGTAATTTGAATTAAC'TTAGAAATTCATTTCTCACCTGTAGT
 GGCCACCTCTATATTGAGGTGCTCAATAAGCAAAAGTGGTGGTGGCTGCTGTATTGGACAGCACAGAAAAAGAT
 TTCCATCACACAGAAAGTGGCTGGCAGCACTGGCCAAGGTGATGGGTGTGCTACACAGTGTATGTCACTGT
 GTAGTGGATGGAGTTTACTGTTTGTGGAATAAAAAACGGCTGTTTCCGTGGAAAAA

235/615

FIGURE 234

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHYRLFPKALGQLISK
YSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWKELSNVLSGIFCASLNFIDSTNTVTPT
ASFKPLGLANDTDHYFLRYAVLPREVVC TENLTPWKLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNA
RCTSISWELRQTL SVVFD AFITGQGKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTY
QDVILGTRKTYAIYDLLDTAMINNSRNLNIQLKWKRPPE NEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTHPYR
AFPVLLLDTPWPYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANSVTKVSIQFERALLKWT
EYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSDGSNYFVRLYTEPLLVLNLPDPFSMPYNVI
CLTCTVVAVCYGSFYNNLLTRTFHIEEPRTGGLAKRLANLIRRARGVPPL

20510" 98526001

236/615

FIGURE 235

TGACGTCAGAATCACCAATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAGCACCCAGGAGC
CCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGTGGGCTACCCCTGGTGGTGG
TTATGGGGGTCTGCCCCCTGGAGGGCCTTATGGACCACCAGCTGGTGGAGGGCCCTATGGACACCCCAATCCTGG
GATGTTCCCTCTGGAATCCAGGAGGACCATATGGCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCC
AAGTTCCTACGGTGGCCAGCAGCCTGGGCTTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTA
CTCCTGGTTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAA
CTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCAGGCCG
CATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGGAAGAACCCTTCCAGCAGTATGACCG
GGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCCAAATGGGCTACAACCTGAGCCCCCA
GTTACCCAGCTTCTGGTCTCCCGCTACTGCCCCAGCTCTGCCAATCCTGCCATGCAGCTTGACCGCTTCATCCA
GGTGTGCACCCAGCTGCAGGTGCTGACAGAGGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCT
CAGCTTCGAGGACTTCGTCAACATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCAC
CAGGGACCTTTCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCCTGCATCATAGCCACCAAAATAGTGAGG
ACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCTGATGGCCATGAGC
AGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCACCC
TGATGCCAGTGGTGAAGTGTTCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAAC
GAGCCCATTTTCTCCAAAGTGAATCTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCT
GCCACACCCATAAATCCTTGTGTGTTAACTTCTAGCTGCCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTG
GGCATCTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCTTCAGT
CTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTTCATTTGGGGCCAAAAG
TCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

20570"9856001

237/615

FIGURE 236

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYGHPNPGMFPSG
TPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVSDHSGYISMKELKQALVNCNWSS
FNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLL
VSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREKDTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

Downloaded from www.ashg.org

238/615

FIGURE 237

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCCTGCTCCTGGTCTGTCTTCATCTCCCAGGCCCT
CTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAAGTTTCCCAAAACTTCGGGACCAACTTGCCCTCAGCTCGGACA
ACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGCTCTGGACCCCTAGGTCTAATGACTTGGCAAG
GGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATGGCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCC
TCCATCGTGGGGGCTGCCTGCCATGGATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGA
GGACCGCTTGGGGGAAGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGG
CCCTTTGCCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGACTCGGA
GTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACGCCCTCCCTGGTCTCT
CATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTGTGTCTTGGGGAGGTGGAGGCCCTGG
GACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGAATCTGGGGTATCAATAATCAACCCCCAGGTACCAG
CTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAA
TATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCTCTGG
AGTTCTCCGCCCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTCAGTG
GGCTTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCTTGCTGTGTGGGCTCA
ATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCCTGCTCATCTCCAATAAAATA
AAAGCACTTATGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

2025.04.04 09:00

239/615

FIGURE 238

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSNFGTNLPQLGQPSSTGSPNSEHPQPALDPRSNDLARV
PLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPDPWQMMAAAEDRLGEALPEELSYLSSAAALAPGSGP
LPGESSPDATGLSPEASLLHQDSESRRLPRSNLSGAGGKILSQRPPWSLIHRVLPDHPWGTLNPSVSWGGGGPGT
GWGTRPMHPPEGIWGINNQPPGTSWGNNINRYPGGSWGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGV
LRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263, 259-264, 269-
274, 270-275, 280-285, 281-286, 305-310

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

FIGURE 239

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCGCTGTGCCTGCTGTGCCCCGCGCTGTGCGCC
CTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTGGAGCCCTGCGGAGAGCTCAAGCGCCC
AGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTCCCATAGTTGCTGCAAGGAGTGGAGCC**AT**GAGCTGCGTCCT
GGTGGTGTGCATCCCCTTGGGGCTGCTGTTCTTGGTCTGCGGATCCCCAAGGCTACCTCCTGCCCCAACGTCACCTC
CTTAGAGGAGCTGCTCAGCAAAATACAGCAACAACGAGTCTCACTCCGGGTCCGAGAGCCATCCCCAGGGAGGA
CAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGACAGCCTCAGGCCTCCAACATGGAGTACAT
GGTGAGCGCCGGCTCCGGCCGAGAGGCTGGCACCGGGGGTGGGGCCTGGGCCACCAGCCTGCTCTGTTCCCCAG
CCAGCTCTGTTCCCCAGCCAGTGCCTGTGATGGCTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCCGGCTCTG
TTCTGTTTTGTTTGTTTGTTTTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCA
TGCCCTGAAACCT**TAG**ACTCCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAAC'TACAGGCATGC
ACCATTGGTCCCCAGCTAGATATTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGTTGCCAGGGCTGGTCTTGAA
CTCCTAGGCTCAAGCAATCCTCCTGCTCAGCCTCTCAAAGTGCTAGGATTATAGGCATGAGTCAACCCTGTCTGG
CTCTGGCTCTGTTCTTAACATTCTGCCAAAAACAACACACGTGGGTTCCTGTGTCAGAGCCTGCCTCGTTGCCTTC
ATGTCAC'TCTTGGTAGCTCCACTGGGAACACAGCTCTCAGCCTTTCCACCTTGAGGCAGAGTGGGGAGGGGCC
AGGGCTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTCCTTAGCCCG
TGTGAGCCTCACTTTCCACTTGGAGAGTCCCTCCTCGCTGGTTGCCATGACTGTGAGATAAGTCGAGGCTGTGA
AGGGCCCGCAGACAGCTGACCTGCCTCCCCAACCCCTAGGCTTTGCTAACCCGGAAAGGAGCTAACCGTGACAGA
AGACAGCCAAGGTC AACCTCCCGGGTGATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGCAC
CCCAAGCTCCAGTGTGGAACCTTCCTTCCTGGCTGGTTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAG
GGTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTGAGCTATGAATGGCTTTTTTAAACAAACC
CACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAATCCAAGTTAGCCGGGCATGGTGGTGCG
CACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAGGTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGC
TTGAGCTTGGGAAGTCGAGGCTGCAGTGAGCTGAGATTGCACCACTGCACCTCAGCCTGGGTGACAGAGCAAGAC
CCTGTCTCAAAAAA

241/615

FIGURE 240

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILMLHNKLRGQVQPQAS
NMEYMSVAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGLRVSSGRGGSRLCSVLFVCFETGSHSATDAGVQ
WHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

2005-05-10 10:50:00

242/615

FIGURE 241

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATG**GGGTCTGGGCTGCCC
CTTGTCTCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGACTTTGCAACTGAAGCTGAAG
GAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCAT
CTCCCTTCAGGGACCAGCGTCACCCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACAT**TGA**CAGCCA
TTGAAGCCTGTGTCCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAG
CAGGCCCCCACCCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

2025-11-19 14:52:50

243/615

FIGURE 242

MGSGLPVLLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTLHHARSQHHVVCNT

20510 9852500T

244/615

FIGURE 243

GGCAAGTGGAACCACTGGCTTGGTGGATTTTGCTAGATTTTCTGATTTTTAACTCCTGAAAAATATCCCAGAT
AACTGTCA**ATGA**AGCTGGTAACTATCTTCCTGCTGGTGACCATCAGCCTTTGTAGTTACTCTGCTACTGCCTTCCT
CATCAACAAAGTGCCCCCTTCCTGTTGACAAGTTGGCACCTTTACCTCTGGACAACATTCTTCCCTTTATGGATCC
ATTAAAGCTTCTTCTGAAAACTCTGGGCATTTCTGTTGAGCACCTTGTGGAGGGGCTAAGGAAGTGTGTAAATGA
GCTGGGACCAGAGGCTTCTGAAGCTGTGAAGAACTGCTGGAGGCGCTATCACACTTGGTGT**TGA**CATCAAGATAA
AGAGCGGAGGTGGATGGGGATGGAAGATGATGCTCCTATCCTCCCTGCCTGAAACCTGTTCTACCAATTATAGAT
CAAATGCCCTAAAATGTAGTGACCCGTGAAAAGGACAAATAAAGCAATGAATACATTA

205110" 9852500T

245/615

FIGURE 244

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59855
<subunit 1 of 1, 93 aa, 1 stop
<MW: 10161, pI: 7.39, NX(S/T): 0
MKLVITIFLLVTISLCSYSATAFLINKVPLPVDKLAFLPLDNILPFMDPLKLLKLTGLISV
EHLVEGLRKCVCNELGPEASEAVKKLLLEALSHLV
```

Important features:

Signal peptide:

Amino acids 1-18

[illegible]

246/615

FIGURE 245

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGATCCCTCTATGACT
GCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGC**ATGA**AGCTCTTATCTTTGGTGGCTGTGGT
CGGGTGTCTTGGTGGTGGCCCCAGCTGAAGCCAACAAGAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACC
TTATAGAAACATCAGTGGGCACATTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGGA
GCCCCATGCCAGTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAGCAC
CACCACCATCAAGGTCATCATTGTCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCTACATGGCCTTCCTGAT
GCTGGTGGACCTCTGATCCGAAAGCCGGATGCATACACTGAGCAACTGCACAATGAGGAGGAGAATGAGGATGC
TCGCTCTATGGCAGCAGCTGCTGCATCCCTCGGGGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGC
CCAGCAGCGGTGGAAGCTGCAGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGCT**TAGAT**
GGGCTGGTGTGGTTGGGTCAAGGCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGGCTACTT
CTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTCTTCTCCTTCTCCCTAACTTTAGAAATG
TTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCTGATCTCTGTTGTCTTCTTGGGTCTTTGGGGTT
GAAGGGAGGGGAAGGCAGGCCAGAAGGGAATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCCATCTGTCTC
TCCTGGCTCCACTCTTGCCGCCTTCCAGCTCTGAGTCTTGGAATGTTGTTACCCTTGGAAGATAAAGCTGGGTC
TTCAGGAACTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTAGCATGTGTTCCCTTCTGCAGTGGTTCTTATC
ACCACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTGAGGACAGCTCTGATGGGAGAGC
TGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCCTGGAAGCTGGTGTTCGCTGTCCCTGTGCACTTCTCGCA
CTGGGGCATGGAGTGCCCATGCATACTCTGCTGCCGGTCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTC
CTCTCCCCAGTGTCCACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGA
ACACCACAGCCCCTGTACTTGGGTTGCCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGAGAGAAAATTTTG
TCCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAAATTGTTTTATTTCTCTCA

405536 01109

247/615

FIGURE 246

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278
<subunit 1 of 1, 183 aa, 1 stop
<MW: 20574, pI: 6.60, NX(S/T): 3
MKLLSLVAVVGCLLVPPAEANKSSDIRCKICPPYRNISGHIYNQNVSQKDCNCLHVVEPMPVPGHDVEAYCLL
CECRYEERSTTTIKVIIIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTEQLHNEEENEDARSMAAAAASLGGPRA
NTVLERVEGAQQRWKLQVQEQRKTTFDRHKMLS
```

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 90-112

N-glycosylation sites.

amino acids 21-24, 38-41 and 47-50

247/615

248/615

FIGURE 247

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAAACTTTCTGAGTTTCAAAAACAACAGA
CTAGTACTCTAAAGAACTCTTTAAAACAATTAACGTTAGGATTGCAGTTATGATTGGATATTATTTAATTCTGT
TTCTGATGTGGGGTTCCCTCCACTGTGTTCTGTGTGCTATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTG
AAAATGAATGCTTAGTGGATCTGTGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAG
TTCCCCCTCCCCCTCCGATTGTTCTAAATTAATTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTGTAT
GATTCTCAACCATCTTTAGTTGGGAAAGGTCCCTTGAAAGCCAATGGAAATACTTTTTTTTTTTCTTGCCACTAAT
CAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTAGTAAAATAGAAACCTGTGTTTATTCTCAG
GTATTTTAGAAACAACAGCCATCATTTTATTTTATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGG
GCTATCAAATATTACTTCATTCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCA
TTTTCTCAGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTATCTTCA
GATTACTTGATTCAAATAAACCAATTATGTTTGTAAATTGATATTAATAAAACCAGAATAAAAGTTCATATCTACCC

2025-10-29 10:43:00

249/615

FIGURE 248

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLPSDCSK

Important features:

Signal peptide:
amino acids 1-29

[illegible]

[illegible]

251/615

FIGURE 250

MAALMRSDSSCCLLLLAAVLMVESSQIGSSRAKLNSIKSSLGGETPGQAANRSAGMYQGLAFGGSKKGKNLGQA
YPCSSDKECEVGRYCHSPHQSSACMVCRRKKRCHRDGMCCPSTRCNGICIPVTESILTPHIPALDGTRHRDR
NHGHYSNHDLGWQNLGRPHTKMSHIKGHEGDPCLRSSDCIEGFCCARHFWTKICKPVLHQGEVCTKQRKKGSHGL
EIFQRCDCAKGLSCKVWKDATYSSKARLHVCQKI

Signal peptide:
amino acids 1-25

20250909 10:00:00

252/615

FIGURE 251

TCTCAATCTGCTGACCTCGTGATCCGCGCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAAGTGTTGGGATTAC
AGGCGTGAGCCACCGCGCCCCGGCCAACATCACGTTTTTAAAAAATTGATTTCTTCAAATTCATGGCAAATATTTCC
CTTCCCTTTAACTTCTTATGTGAGAATGAGGAAGGATAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAAT
ATTTTCATGTAGTATTTTCTAAGTTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTG
AAAATACTTGATGTGTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTATCCCCCTTTT
AAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGACCACCAAAAATGGCAAACATCACCAGCTCCCAGAT
TTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACCCCAAGTACACAGCAGAATAGTACAAGTCA
CCCTACAATACTACTTCTTGGGACCTCAAGCCCCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAA
ATCTCAACCTGAGCCATCCCCAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACGTGT
TCCTCCTCCTGGTTTGGAGTCCTTTCTTCCCAGGCAAAACTTCGAGAATCAACACCTGGAGACAGTCCCTCCAC
TGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTTGAAAATATCTCTGTGTCTGTCCACCAGCCACAGCCCAA
ACACATCAAATTTGCTAAGCGGCGGATACCCCCAGCTTCTAAGATCCCAGCTTCTGCAGTGGAATGCCTGGTTC
AGCAGATGTCACAGGATTAAATGTGCAGTTTGGGGCTCTGGAATTTGGGTGAGAACCCTTCTCTCTGAATTTGG
ATCAGCTCCAAGCAGTGAAAATAGTAATCAGATTCCCATCAGCTTGTATTGGAAGTCTTTAAGTGAGCCTTTGAA
TACATCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACCTCCGTCATTACCTCCTGCAGTCTGAC
AAGCTCATCACTGAATTTCTGCTAGTCCAGTAGCAATGTCTTCTCTTATGACCAGAGTTCTGTGCATAACAGGAT
CCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCAGGAACCATCATGAATGGACATGGTGGTGGTCTGAAG
TCAGCAGACACTAGACAGTAAGTATAGCAGCAAGCTACTCTTGTCTATGGCTGGTGCCAACCAAACAGAGGAAGAG
GATAGCTCACGTGATGTGGAAAACACCAGTTGGTCAATGGCTCATTTCGTTAAAGCAGCCCTTTTGCTTTTTTTG
TTTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGTGGTGGTCTCATA
TTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATAAGTATGCATTTTAAAGATGCTTGGGCCAGGCGGG
GTGGCTGATGCCCATAAATCCCAGTGCTTTGGGGGGCCAAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAGACC
ACCCTGGGCAACATGGTGAAGCTCTGTCTCTACTAAAAATACGAAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCC
TGTAATCCCAGCTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCCGT
CTGAAAAGA

20510 "9955001

253/615

FIGURE 252

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTTPSTQQNSTSHPTT
TTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAVTVPPPGLESFPSQAKLRESTPGDSPSTVNK
LLQLPSTTIENISVSVHQPPKHIKLAKRRIPPASKIPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAP
SSENSNQIPISLYSKSLSEPLNTLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPIYQ
SPVSSSESAPGTIMNGHGGGRSQQTLDISKYSSKLLLSWLVPKQQRKRIAHVMWKTPVGQWLIR

Signal peptide:
amino acids 1-24

253/615

[illegible]

255/615

FIGURE 254

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQKQYQIRIRKEKPQQH
NFTHRLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQREHRSMRANVELDHATLVRFSPCRA
FIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPPKKHKAPVIDIGIANTGKFIMTASSD'TTVLIWSLKGQVLSTI
NTNQMNNTTHAAVSPCGRFVASCFTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKD
GTWKLWDTDVEYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVHGEC
IANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQLQQQLTQAQETLKS LGALKK

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

2025-04-04 10:00:00

256/615

FIGURE 255

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGCCAGGTGCCCC
GTCGCAGGTGCCCCCTGGCCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCCCTTCCTCGGCGCTGCCAACCC
GCCACCCAGCCC**ATGG**CGAACCCCGGGCTGGGGCTGCTTCTGGCGCTGGGCCTGCCGTTCCCTGCTGGCCCCGCTGG
GGCCGAGCCTGGGGGCAAATACAGACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCC
AGCTCCGATGGCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTC
CTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCACT
AGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCTGCCTG
CCCATC**TAGG**TCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCT
GGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCCTGAGGTCAAGAGA
GGATGGGGCTATTCACCTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

2025-09-25 10:00:00

257/615

FIGURE 256

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFSLLAALLLAVG
LALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

259/615

FIGURE 258

MGLFRGFVFLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFEATEKRFFFKN
VSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEYIHFTPDLLLGGKKQNEYGPPG
KLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQF
FPDKVQTEKASIMFMSIDSVVEFCNEKTHNQEAPSLQNIKCNRSTWEVISNSEDFKNTIPMVTTPPPPPVFSLL
KISQRIVCLVLDKSGSMGGKDRLNRMNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGL
PTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMS
KITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFLITWNS
LPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKM
NKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNMGVYSRYFTAYTENGRYS
LKVRAHGGANTARLKLRLPLNRAAYIPGWVNGEIEANPPRPEIDEDTQTTLLEDFSRASGGAFVVSQVPSLPLP
DQYPPSQITDLDATVHEDKIILTWTAPGDNFVDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKES
FAFKPENISEENATHIFIAIKSIDKSNLTSKVSANIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTL
VLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632,
811-815, 832-836, 837-841, 852-856, 896-900

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

260/615

FIGURE 259

CGCCGGAGGCAGCGGCGGCGTGGCGCAGCGGCGACATGGCCGTTGTCTCAGAGGACGACTTTTCAGCACAGTTCAA
ACTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGCACTGCTTGAGAAGCTGCTGGACCGCC
CGCCCCCTGGCCTGCAGAGGCCCCGAGGACCGCTTCTGTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCA
GTCTACTGCCATGGAACCTTCTTTATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAG
CCACCGGGGAGGACCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCGTGC
CCTCCATGCTGTGCCTGGTGGCCAACCTTCTGTCTGTCAACAGGGTTGCAGTCCACATCCGTGTCTCTGGCCTCAC
TGACGGTCACTCTGGCCATCTTCATGGTGATAACTGCACCTGGTGAAGGTGGACACTTCTCTCTGGACCCGTGGTT
TTTTTGCGGTCACCATTTGTCTGCATGGTGATCCTCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCA
TGACCGGCTCCTTTCTATGAGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCG
TGGCCTCATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCTGACGGCCACCA
TCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCCAGGTACTACATGAGGCCTG
TTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCAGGACTCCCTCAGTGCCCCCTTCGGTGGCCTCCA
GATTCAATTGATTCCACACACCCCCCTCTCCGCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCT
ACGTCTTCTTCATCACCAGCCTCATCTACCCCGCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCT
CACTGTGGACCACCAAGTTTTTTCATCCCCCTCACTACCTTCTCTCTGTACAACCTTTGCTGACCTATGTGGCCGGC
AGCTCACCGCCTGGATCCAGGTGCCAGGGCCCCAACAGCAAGGCGCTCCAGGGTTCTGTCTCTCCGGACCTGCC
TCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGTCCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGT
ACCCCGCACTCCTCAGCTCCCTGCTGGGGCTCAGCAACGGCTACCTCAGCACCCCTGGCCCTCCTCTACGGGCCTA
AGATTGTGCCCAGGGAGCTGGCTGAGGCCACGGGAGTGGTGATGTCTTTTATGTGTGCTTGGGCTTAACACTGG
GCTCAGCCTGCTCTACCTCCTGGTGACCTCATCTAGAGGGAGGACACAAGGACATTGGTGCTTCAGAGCCTT
TGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCCTAAAGTTTCACTTGGGGACAGAGAG
CAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCAGTGAGCCACGTCCATGCCCATTCCGTGCAAGGCAGATA
TTCCAGTCATATTAACAGAACACTCCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACA
GCTGATGGTTAAACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTCGCCCTAGAGTTATTACAAA
GCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCTGACAGCGAGATGCAA
GCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGAAGTCCCTGGCATGGTCAGTCCCTCAGGC
CCAAGACTCAAGTGTGCACAGACCCCTGTGTTCTGCGGGTGAACAACCTGCCCCTAACCAGACTGGAAGCCAG
AAAGATGGGCCTTCCATGAATGCTTCATTCCAGAGGGACCAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCT
GGCCTGGGTTTTTCAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTCTTTCA
GTGTTCTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAAATACTTGCGTATTCAAAAA

CGCTTGGGCTTAACACTGG

261/615

FIGURE 260

MAVVSEDDFQHSSNSTYGTTSSSLRADQEALLEKLLDRPPPGLQRPEDRFCGTYYIFFSLGIGSLLPWNFFITAK
EYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFLLVNRVAVHIRVLASLTVILAI FMVIT
ALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSIYGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSD
VRNSALAFFLTATIFLVLCMGLYLLLSRLEYARYYMRPVLAHVFSGEEELPQDSLAPSVASRFIDSHTPPLRP
ILKKTASLGFCVTYVFFITS LIYPAVCTNIESLNKSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVPGPN
SKALPGFVLLRTCLIPLFVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLALLYGPKIVPRELAEATG
VMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252, 305-330,
448-472

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 261

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAAACCACCACACCTGTTTA
 AAGAACCTAAGCACCATTAAAGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATAAAGGAGGGCAGAAT
GGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTACGTGGCCGGAATCATTCCCTT
 GGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGACTGTTTTGGGTGCTGGCCTTCTCTGTGGAACCTGCTCT
 GGCAGTCATCGTGCCTGAAGGAGTACATGCCCTTTATGAAGATATTCTTGAGGGAAAACACCACCAAGCAAGTGA
 AACACATAATGTGATTGCATCAGACAAAGCAGCAGAAAAATCAGTTGTCCATGAACATGAGCACAGCCACGACCA
 CACACAGCTGCATGCCTATATTGGTGTTCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGG
 TAACTCCCATGTGCATTCTACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCAAAATCACCACCACGCTGGG
 TCTGGTTGTCCATGCTGCAGCTGATGGTGTGGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTT
 AATTGTGTTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTTCTTCTTGATGCATGCTGG
 CTTAGAGCGGAATCGAATCAGAAAGCACTTGCTGGTCTTTGCATTGGCAGCACCAGTTATGTCCATGGTGACATA
 CTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTTCAGAGGTGAACGCCACGGGAGTGGCCATGCTTTTTCTCTGC
 CGGGACATTTCTTTATGTTGCCACAGTACATGTCTCCTCGAGGTGGGCGGAATAGGGCAGCCACAAGCCCGA
 TGCCACGGGAGGGAGAGGCCCTCAGCCGCCCTGGAAGTGGCAGCCCTGGTTCTGGGTTCGCTCATCCCTCTCATCCT
 GTCAGTAGGACACCAGCATTAATATGTTCAAGGTCCAGCCTTGGTCCAGGGCCGTTTGCCATCCAGTGAGAACAGC
 CGGCACGTGACAGCTACTCACTTCCTCAGTCTCTTGTCTCACCTTGCGCATCTCTACATGTATTCCCTAGAGTCCG
 GAGGGGAGGTGAGGTAAACCTGAGTAATGGAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTA
 TAGACATCCCATTGTGTTATCTTTTAAAGGCCCTTGACATTTTGCCTTTTAAATATTTCTCTTAACCTATTCTC
 AGGGAAGATGGAATTTAGTTTTTAAAGGAAAAGAGGAGAACCTCATACTCACAATGAAATAGTGATTATGAAAATAC
 AGTGTCTGTAAATTAAGCTATGTCTCTTCTCTTAGTTAGAGGCTCTGCTACTTTATCCATTGATTTTTTAACA
 TGGTTCCCACCATTGTAAGACTGGTGCTTTAGCATCTATGCCACATGCGTTGATGGAAGGTGATAGCACCACCTCA
 CTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGCAAGACACATTGAAAGCTC
 TCTTTATACTCAAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATTAAACAGCTCCTTTGGCACGTGCCT
 CTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGGTGGGAGGAGCTTCTAAAGAGGTGACTGGTATT
 TTGTAGCATTCCTTGTCAAGTTCCTCTTTCGAGAATACCTGTCTCCACATTCCTAGAGAGGAGCCAAGTTCCTAGT
 AGTTTCAGTTCCTAGGCTTTCTTCAAGAACAGTCAGATCACAAAGTGTCTTTGGAAATTAAGGGATATTAAATTT
 TAAGTGATTTTTGGATGGTTATTGATATCTTTGTAGTAGCTTTTTTTAAAGACTACCAAATGTATGGTTGTCC
 TTTTTTTTTGTTTTTTTTTTTTTTTAAATTATTTCTTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGG
 TCAGCTTTGGCGACACTGTGTCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTTCGCT
 ATTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCTTTTCTTTGCTTTTCTTCTAACTTTTCCCTC
 TAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTAAATATTTGTGTGGGATGAATTCCTTATCAGGACAA
 CCACCTCTCGAACTGTAATAATGAAGATAATAATATCTTTATTCTTTATCCCCCTTCAAAGAAATTACCTTTGTG
 TCAAATGCCGCTTTGTTGAGCCCTTAAATATACCACCTCCTCATGTGTAAATTGACACAATCACTAATCTGGTAAT
 TTAAACAATTGAGATAGCAAAAGTGTTTAAACAGACTAGGATAATTTTTTTTTTTCATATTTGCCAAAATTTTTGTAA
 ACCCTGTCTTGTCAAATAAGTGTATAATATTGTATTATTAATTTATTTTTTACTTTCTATACCATTTCAAAACACA
 TTACACTAAGGGGGAACCAAGACTAGTTTCTTCAGGGCAGTGGACGTAGTAGTTTGTAAAAACGTTTCTATGAC
 GCATAAGCTAGCATGCCTATGATTTATTTCTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTT
 GTGAGCCCTCTGCTGGCCACAGTGAGGAAAGTAGCACAATAGGATACAGTTGTATGTAGTCATTGGCAACAATT
 GCATACAATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAAATGACTTCCTTGATTGGATGTTAACAGCT
 GACTGGTGTGAGACTTGAGGTTTCATCTAGTCTTCAAACATATATGGTTGCCTAGATTCTCTCTGGAAACTGAC
 TTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAAA

263/615

FIGURE 262

MDDFISISLLSLAMLVGCVVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHALYEDILEGKHHQAS
ETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQIGNSHVHSTDDPEARSSNSKITTTTL
GLVVHAAADGVALGAAASTSQTSVQLIVFVAIMLHKAPAAFGLVSFLMHAGLERNRIRKHLVLFALAAPVMSMVT
YLGLSKSSKEALSEVNATGVAMLF SAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLI
LSVGHQH

Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 37-56, 106-122, 211-230, 240-260, 288-304

263/615

FIGURE 263

CTCCTTAGGTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGG
GCAGGGGTGACAACAGGTGTCTATCTTTTTGATCTCGTGTGTGGCTGCCTTCCCTATTTCAAGGAAAGAC
GCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCC
AGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGG
TTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGAGGGTG
GACCAGAAGAAAGGAAAGGTCCCTCTTGTCTGTGGCTGCACATCAGGAAGGCTGTGATGGGAATGAA
GGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAAAGTAGAGA
AGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGC
TCCGGGGGGCCCCAAACGCATGCTTCCCTGTGGTCTAGCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCT
TTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAAATGATGATGGTTTCGCCGGGGCTGCT
TGCGTGGATTTCCCGGGTGGTGGTTTGTCTGGTGTCTCTGTCTGTCTATCTCTGTCTGTACATGT
TGGCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAG
GAGGGTACCAGGCTCCTTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGAGCAGCTGAGCGC
GCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAATGGGCAGTACCAAG
CCAGCGATGCTGTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTCCTGGCC
TTCCTGCACCTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGC
AGTGCCTTTTCGATAGCTTTTACTCTACAGAAGGTGTACAGCTGGAGACTGGCCTTACCCGCCACCCCG
AGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACC
CTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCTTACACGGCCTCTGATTTTATAGAAGGGAT
CTACCGAACAGAAAGGACAAGGGACATTGTATGAGCTCACCTTCAAAGGGGACCACAAACAGCAAT
TCAAACGGCTCATCTTATTTTCGACCATTACAGCCCCATCATGAAAGTGAAAAATGAAAAGCTCAACATG
GCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTCGGCAGTTTATGCA
GAATTTTCAAGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGGAAAG
AAGAAATAAATGAAGTCAAAGGAATACTTGA AAAACACTTCAAAGCTGCCAACTTCAGGAACCTTAC
TTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGACTTGATGTTGGAGCCCGCTTCTGGAAGGG
AAGCAACGTCTCTTCTTTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTA
GGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTTCTTTTTCAGTCACTACAATCCTGGCATA
ATATACGGCCACCATGATGCAGTCCCTCCCTTGAACAGCAGCTGGTCATAAAGAAGGAAACTGGATT
TTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCGGTGAGACTTCATCAATATAGGTGGGTTTG
ATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTCCACAGCAACCTC
ATAGTGGTACGGACGCCGTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAGCT
GACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGG
GCATGCTGGTGTTCAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAAAGACAAGTAGCAAA
AAAACATGAAGTCCCAGAGAAGGATTGTGGGAGACACTTTTTCTTTCTTTTGCAATTACTGAAAGTG
GCTGCAACAGAGAAAAGACTTCCATAAAGGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAA
ATACCTCCGATTTCTCTGTGGTGGCTTTTTTACAACAGAAATCAAAATCTCCGCTTTGCTGCAAAAGT
AACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTG
TGGAGGTTTTGATGGTGTTTACAATACACTGAGACCTGTTGTTTTGTGTGCTCATTGAAATATTCATG
ATTTAAGAGCAGTTTTGTAAAAAATTCAATAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTTCTAGGAATGCTAAAAATATCAGAAGGCAGGAGAGGATAGGCTTA
TTATGATACTAGTGAGTACATTAAAGTAAAAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCG
TGTCAATTTTTCCCCAAGATTAACCAAAAATAATCTGCTTATCTTTTGGTTGTCTTTTAACTGTCT
CCGTTTTTTTTCTTTTATTTAAAAATGCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTA
CCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTACATTTTTATATTTTTTAAGAAGATACTTT
GAGATGCATTATGAGAACTTTCACTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATG
CTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATAACAGA
CGTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATG
ACACTTTCTGCTTTTACAGAAAAGGAAACTCAATTCAGACTGGTGATATCGTGATGTACCTAAAAGTCAG
AAACCACATTTTCTCTCAGAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAGTATACCGT
GTGAACCAACAATCTCTTTTCAAAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATG
GAGAAAAATATATATATATATATATATATTTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATG
GAAGTTTTTGTCTACATGTTATCCACCCCAGGCCAGGTGGAAGTAACTGAATTATTTTTTAAATTAAGC
AGTTCTACTCAATCACCAAGATGCTTCTGAAAATTGCATTTTTATTACCATTTCAAACATTTTTTTAAA
AATAAATACAGTTAAATAGAGTGGTTTCTTCAATTCATGTGAAAATTATTAGCCAGCACCCAGATGCAT
GAGCTAATTTATCTCTTGTAGTCCTTGTCTGTTGTGCTACAGTAACTCATTGTTTAAAGCTTCAA
GAACATTCAAGCTGTTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTT
AATTAACACAGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGGATATGAA

265/615

FIGURE 264

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQALALPRANSPTGKEGYQAVLQEWEEQHRNYVSSL
KRQIAQLKEELQERSEQLRNGQYQASDAAGLGGLDRSPPEKTQADLLAFLHSQVDKAEVNAGVKLATEYAAVPFDS
FTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYE
LTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMANLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLLTVVY
FGKEEINEVKGILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNT
QPGKKVFYPVLFSEQYNPGIYYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGED
VHLYRKYLHSNLIVVRTPVRLFHLWHEKRCMDLTPQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHLRKQKQ
KTSSKKT

2025-09-24 10:00:00

FIGURE 265

GGATG CAGAAAGCCTCAGTGTGCTCTTCTGGCCTGGGTCTGCTTCTCTTCTACGCTGGCATTGCCCTCTTCA
 CCAGTGGCTTCTGCTCACCCGTTTGGAGCTACCAACCATAGCAGCTGCCAAGAGCCCCAGGCCCTGGGTCCC
 TGCCATGGGGGAGCCAAGGGAAACCTGGGGCCTGCTGGATGGCTTCCCGATTTTCGCGGGTGTGTTGGTGCTGA
 TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT
 TCCTGGGCAAACCTAAGCTCCTTGCAGAGGATCCTGGAGATTACAGCCCCACCATGCCCGCTCTACCGATCTCAGG
 TTGACCCTCCTACCACCACCATGCAGCGCCTCAAGGCCCTCACCAGTGGCTCACTGCCTACCTTTATTGATGCTG
 GTAGTAACCTTCGCCAGCCACGCCATAGTGGAAAGACAATCTCATTAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG
 TCTTCATGGGAGATGATACCTGGAAAGACCTTTTCCCTGGTGCTTTCTCCAAAGCTTTCTTCTCCCATCCTTCA
 ATGTGAGAGACCTAGACACAGTGGACAATGGCATCCTGGAACACCTCTACCCACCATGGACAGTGGTGAATGGG
 ACGTGTGATTGCTCACTTCCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA
 AGAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG
 CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC
 TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCACAGAGGCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
 CCACGCTGGCCCTGCTGCTGGGCCTGCCCATTCCATTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG
 GGGGTGAGGACTCCCAGCCCCACTCCTCTGCTTTAGCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
 CCCGATTTCTTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCAGAACCTCT
 TCTCCAAGGCCTCTGCTGACTACCACTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA
 TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC
 GCATGGCGGGGGGTACTGCTCTCTTGGCTGCTTCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC
 CAGGCTTTCCATTCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
 TCCTGGGAACCTATTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTATTCTCCCTT
 TTCTGTGGAAGCCTGGGCTGGCTGGGGGTCCAAGAGGCCCTGGCAACCCTGTTTCCCATCCCTGGGCCCCGTCC
 TGTACTCCTGCTGTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTGTGTGAGCTGAGGCCAGGGCCACCCCT
 TCCTTTTGGGCTCATTATCCTGCTCCTGGTTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC
 TCACAATGCCCCGCCTTGGCACTTCAGCCACAACAACCCCCACGGCACAAATGGTGATATGCCCTGAGGCTTG
 GAATTGGGTTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTCATCGTTGCCCTGAAGAGACACCTGTTTGGCACT
 CCTCTCCCTGGCTGAGTCTCTGGCATCCATGGTGGGTGGTGCAGCCAAGAATTTATGGTATGGAGCTTGTGTGG
 CGGCGCTGGTGGCCCTGTTAGCTGCCGTGCGCTTGTGGCTTCGCCGTATGGTAATCTCAAGAGCCCCGAGCCAC
 CCATGCTCTTTGTGCGCTGGGGACTGCCCTAATGGCATTGGGTAAGTGTGCTGCTGCTACTGGGCATTGGCGTGGGGG
 CAGATGAGGCTCCCCCGCTCTCCGGTCTGGTCTCTGGGGCATCCATGGTGCTGCCCTCGGGCTGTAGCAGGGC
 TGGCTGCTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGTGGTGAAGGCTGGGGCAGGCGCTCCAA
 GGACCAGGACTGTCTCACTCCCTTCTCAGGCCCCCCCCACTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAA
 TCTACCGACACATGCAGGAGGAGTTCCGGGGCCGTTAGAGAGGACCAAATCTCAGGGTCCCCTGACTGTGGCTG
 CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCAAGCCCTCACCCTGTTGGCCTTCCCACTTCTGCTGT
 TGCATGCGGAGCGCATCAGCCTTGTGTTCTGCTTCTGTTTCTGCAGAGCTTCTTCTCTCTACATCTGCTTGCTG
 CTGGGATACCCGTCACCACCCCTGGTCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCCATGGCCA
 CACAGACCTTCTACTCCACAGGCCACAGCCTGTCTTCCAGCCATCCATTGGCATGCAGCCTTCGTGGGATTCC
 CAGAGGGTCTAGGCTCCTGTACTTGGCTGCCCTGCTTTGCTAGTGGGAGCCAACACCTTTGCCCTCCACCTCCTCT
 TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTCTCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC
 CCCCAGGGAATGAAGCTGATGCCAGAGTCAGACCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG
 ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCTCAAGTACCTCTTTATCCTTGGTATTCAGATTCT
 TGGCCTGTGCCCTTGGCAGCCTCCATCCTTCGCAGGCATCTCATGGTCTGGAAGTGTGGCCCCAAGTTTCATAT
 TTGAGGCTGTGGGCTTCAATTGTGAGCAGCGTGGGACTTCTCTGGGCATAGCTTTGGTGATGAGAGTGGATGGTG
 CTGTGAGCTCCTGGTTCAGGCAGCTATTTCTGGCCACAGCAGGTTAGCCTAGTCTGTGATTACTGGCACTTGGCT
 ACAGAGAGTGTGGAGAACAGTGTAGCCTGGCCTGTACAGGTAAGTGGATGATCTGCAAGACAGGCTCAGCCATAC
 TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCATTATTCTATAATTAGGACCACAGTGGAGTA
 TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGACAAGGGGGCGGTCTCCGAAGTGGAAATAAATAGGCCGG
 GCGTGGTGACTTGCACCTATAATCCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTTGGTCCCAGGAGTTCA
 AGACCAGCCTGTGGAACATAACAAGACCCGCTCTACTATTTAAAAAAGTGAATAAAATGATAATAT

267/615

FIGURE 266

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPGPGSLPWGSQKPGACWMASRFSRVVLVLI
DALRFDFAQPPQHSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQVDPPTTTMQRLKALT'TGSLPTFIDAG
SNFASHAIVEDNLIKQLTSAGRRVVFMDGDDTWKDLFPGAFSKAFFFPFNFVRDLDTVDNGILEHLYPTMDSGEWD
VLIAHFLGVDHCGHKHGHHPPEMAKKLSQMDQVIQGLVERLENDTLLVVAGDHGTTNGDHGGDSELEVSAALFL
YSPTAVFPSTPPEEPEVIPQVSLVPTLALLLGLPIPFNGNIGEVMAELFSGGEDSQPHSSALAAQASALHLNAQQVS
RFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRGARAMCIESWARFSLVR
MAGGTALLAASCFICLLASQWAISPGFPFCPLLLTPVAWGLVGAIAYAGLLGTIELKLDLVLLGAVAAVSSFLPF
LWKAWAGWGSKRPLATLFPPIPGPVLLLLLLFRLAVFFSDSFVVAEARATPFLGSEFILLLVVQLHWEGQLLPPKLL
TMPRLGTSATTNPPRHNGAYALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVA
ALVALLAAVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPRLRVLVSGASMVLPRVAGL
AASGLALLLWKPVTVLVKAGAGAPRTRTVLTPTFSGPPTSQADLDYVVPQIYRHMQEEFRGRLETKSQGPLTVAA
YQLGSVYSAAMVTALTLAFPLLLLHAERISLVFLLLFLQSFLLLHLLAAGIPVTTTPGPFTVPWQAVSAWALMAT
QTFYSTGHQPVFPAIHWHAFAVGFPEGHGSCTWLPALLVGANTFASHLLFAVGCPLLLLPFLCESQGLRKRQQP
PGNEADARVRPEEEEEPLMEMRLRDAPQH FYAALLQLGLKYLFI LGIQLACALAA SILRRHLMVWKVFAPKFIF
EAVGFIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850, 1016-1034,
1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

[illegible]

Author	Year	Country	Sample Size	Study Design	Findings
Wang et al.	2000	China	1,000	Case-control	Increased risk of lung cancer with tobacco use.
Li et al.	2001	China	2,000	Cohort	Increased risk of lung cancer with tobacco use.
Chen et al.	2002	China	1,500	Case-control	Increased risk of lung cancer with tobacco use.
Yang et al.	2003	China	3,000	Cohort	Increased risk of lung cancer with tobacco use.
Zhang et al.	2004	China	1,200	Case-control	Increased risk of lung cancer with tobacco use.
Wu et al.	2005	China	2,500	Cohort	Increased risk of lung cancer with tobacco use.
Qin et al.	2006	China	1,800	Case-control	Increased risk of lung cancer with tobacco use.
He et al.	2007	China	2,200	Cohort	Increased risk of lung cancer with tobacco use.
Guo et al.	2008	China	1,600	Case-control	Increased risk of lung cancer with tobacco use.
Liu et al.	2009	China	2,800	Cohort	Increased risk of lung cancer with tobacco use.
Wang et al.	2010	China	1,400	Case-control	Increased risk of lung cancer with tobacco use.
Li et al.	2011	China	2,100	Cohort	Increased risk of lung cancer with tobacco use.
Chen et al.	2012	China	1,900	Case-control	Increased risk of lung cancer with tobacco use.
Yang et al.	2013	China	3,100	Cohort	Increased risk of lung cancer with tobacco use.
Zhang et al.	2014	China	1,300	Case-control	Increased risk of lung cancer with tobacco use.
Wu et al.	2015	China	2,600	Cohort	Increased risk of lung cancer with tobacco use.
Qin et al.	2016	China	1,700	Case-control	Increased risk of lung cancer with tobacco use.
He et al.	2017	China	2,300	Cohort	Increased risk of lung cancer with tobacco use.
Guo et al.	2018	China	1,500	Case-control	Increased risk of lung cancer with tobacco use.
Liu et al.	2019	China	2,900	Cohort	Increased risk of lung cancer with tobacco use.
Wang et al.	2020	China	1,400	Case-control	Increased risk of lung cancer with tobacco use.

FIGURE 268

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812
><subunit 1 of 1, 337 aa, 1 stop
><MW: 37668, pI: 6.27, NX(S/T): 1
MLFSAILLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQPHGYDQPGTE
PLDLHNNNGHTVQLSLPSTLYLGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEATFAELHIVHYDSDSYDSLSEA
AERPQGLAVLGILIEVGETKNIAIEHILSHLHEVRHKDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSV
LWTVFYRRSQISMEQLEKLKQGLTFSVEEPSKLLVQNYRALQPLNQRMVFASFQAGSSYTTGEMLSLGVGILVG
CLCLLLAVFYFIARKIRKLENRKRVSVFTSAQAATTEA
```

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

270/615

FIGURE 269

GTGGCGCTGGCGGTTGCTGTCTCAGCTGATTCCCGGGGTTGGTGGCAGCGGCGGTAGCAGCAATGGACTTTCTCCTG
GGGAACCCGTTTCAGCTCTCCAGTGGGACAGCGCATCGAGAAAGCCACAGATGGCTCCCTGCAGAGCGAGGACTGG
GCCCTCAACATGGAGATCTGCGACATCATCAACGAGACGGAGGAAGGTCCCAAAGATGCCCTCCGAGCAGTAAAG
AAGAGAATCGTGGGGAATAAGAACTTCCACGAGGTGATGCTGGCTCTCACAGTCTTAGAAACCTGTGTCAAGAAC
TGCGGGCACCGCTTCCACGTGCTGGTGGCCAGCCAGGACTTCGTGGAGAGTGTGCTGGTGAGGACCATCCTGCCC
AAGAACAACCCACCCACCATCGTGCATGACAAAGTGCTCAACCTCATCCAGTCTTGGGCTGACGCGTTCGCGCAGC
TCGCCCCGATCTGACAGGTGTGGTCACCATCTATGAGGACCTGCGGAGGAAAGGCCTGGAGTTCCCCATGACTGAC
CTGGACATGCTGTACCCATCCACACACCCAGAGGACCGTGTTCAACTCAGAGACACAATCAGGACAGGATTCTG
TGGGCAC TGACTCCAGCCAGCAAGAGGACTCTGGCCAGCATGCTGCCCCCTCTGCCCCCCCCGCCCCATACTCTCCG
GTGACACGCCCCATAGCACCAACCCCGGAACAGATTGGGAAGCTGCGCAGTGAGCTGGAGATGGTGAGTGGGAACG
TGAGGGTGATGTCTGGAGATGCTGACGGAGCTGGTGCCACCCAGGCCGAGCCCGCAGACCTGGAGCTGCTGCAGG
AGCTCAACCGCACGTGCCGAGCCATGCAGCAGCGGGTCTGAGTGATACCTGCTCCGGGGCCCATGCCCCAAGGA
GCCCTTCAGAGCCCACACTGCCAGTCGAGGCCCTGGCTGGAGGCTGGCCACAGTGGAATTTCTGCCGAGCCTATTG
TCCCTACCCTGCTCTGCTGCATGGGGCCCCATGGCTTTGGCTGGCCACTGAGGGTAGGGTGTGGAGGTGTGGAGG
CCCCCTGAGGAGCTGCGGCGGCCAGGTACGAAGCTGCAACTCTGCGCGCAGTGGGCGAGATCTCATCAGCCCCA
GGCTGCAGGTGAGGCTTCAGGGGATGCTGGGGCCCCACTGCCCTCCGCTGCCTTGCCCTCCATCCTTCCTCTGT
TCCTTCTGGCCGGGCACACAGCACTGGGGCTCACCTCTTGTTGATCCTCTTGTACTGGGAGAGGTGCCTTTTG
TATCCCCAATTAAAGGTAGAAAACC

270/615

271/615

FIGURE 270

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62813
><subunit 1 of 1, 209 aa, 1 stop
><MW: 23465, pI: 7.57, NX(S/T): 1
MDFLLGNPFSSPVGQRIEKATDGSLSQSEDWALNMEICDIINETEEGPKDALRAVKKRIVGNKNFHEVMLALTVLE
TCVKNCGRHFHVLVASQDFVESVLVRTILPKNNPPTIVHDKVLNLIQSWADAFRSSPDLTGVVVTIYEDLRRKGLE
FPMTDLDMLSPIHTPRGPCSTQRHNQDRILWALTPASKRTLASMLPLCPFRPFYSPVTRP
```

Important features of the protein:

Signal peptide:

Amino acids 1-15

N-glycosylation site:

Amino acids 41-45

N-myristoylation sites:

Amino acids 6-12;23-29

271/615

272/615

FIGURE 271

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCGTCTGCCATGGGGC
TCGGGTTGAGGGGCTGGGGACGTCCTCTGCTGACTGTGGCCACCGCCCTGATGCTGCCCCTGAAGCCCCCGCAG
GCTCCTGGGGGGCCAGATCATCGGGGGCCACGAGGTGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCT
TCGGGGGCCAACATCACTGCGGAGGCTTCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGGCCACTGCTTCAGCC
ACAGAGACCTCCGCACTGGCCTGGTGGTGTGCTGGGCGCCACGTCCTGAGTACTGCGGAGCCCCACCCAGCAGGTGT
TTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGACATCTGCCCTGCTGCGGC
TGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCTGCCAGGGAGAAGGGCCAGGCCCCCACAG
CGGGGACACGGTGCCGGGTGGCTGGCTGGGGCTTCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGG
CCAAGGTCCGAGTGTGCGACCCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCA
CCCCGAGTGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGCCCCCTGGTGTGCAGGAACCGGG
CTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCGACGTGTACACGCAGGTGTCCG
CCTTTGTGGCCTGGATCTGGGACGTGGTTTCGGCGGAGCAGTCCCCAGCCCCGGCCCCCTGCCCTGGGACCACCAGGC
CCCCAGGAGAAGCCGCCTGAGCCACAACCTTGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCG
TGGCTGGGCCCCACGGGAAGCCTGATGTTTCAGGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACA
TGCAAAGGGCAGAAGCAAACCCAGTAAAAATGTTAACTGACAAAAAAAAAAAAAAAAAAAAAGAAA

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCGTCTGCCATG

273/615

FIGURE 272

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845
><subunit 1 of 1, 283 aa, 1 stop
><MW: 30350, pI: 9.66, NX(S/T): 2
MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPYMASVRFGGQHHCGGFLLRARWVVSAAHC
FSHRDLRTGLVVLGAHVLSTAEP TQQVFGIDALTTHPDYHPMTHANDICLLRLNGSAVLGPAVGLLRLPGRRARP
PTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPDVCNSSWKGHLLTLM LCTRSGDSHRRGFCSADSGGPLVCR
NRAHGLVSFSGLWCGDPKTPDVYTQVSAFVAWIWDVVRSSPQPGPLPGTTTRPPGEAA
```

Signal peptide:
amino acids 1-30

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 273

GAAGTTCGCGAGCGCTGGCT**ATG**GGTCCTGGGGCGCGGGCTGGCGGCGCTGCTGGCGGTGCTGGCGCTCGGGACAG
 GAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGACCAGCGTGGCGCGCGCCCTGGCGCCCC
 AGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGCGCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCT
 ACGACAAGGTACTTTCTTTGCATGAGGATTCAACAACCCCTGTGGCTAACCCCTCTGCTTGCAATTTACTCTCATCA
 AACGCCTGCAGTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGGATG
 GCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGGGCCCTGATGCGGCTGC
 AGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCCGAGGTGTCTTTTCAGAGAGTCACTGGCTCTGCCATCACTG
 ACCTGTACAGCCCCAAACGGCTCTTTTCTCTCACAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACA
 TGGGGGATTATTACCATGCCATTCCATGGCTGGAGGAGGCTGTCACTCTCTCCGAGGATCTTACGGAGAGTGGA
 AGACAGAGGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAAATGTTT
 CGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCAGATAATAAGAGGATGGCCAGGAATGTCTTGA
 AATATGAAAGGCTCTTGGCAGAGAGCCCCAACACGTTAGCTGAGGCTGTCTCCAGAGGCCCAATATACCCC
 ACCTGCAGACCAGAGACACCTACGAGGGGCTATGTACAGCCCTGGGTTCCAGCCCCACTCTCTACCAGATCCCTA
 GCCTCTACTGTTTCTATGAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCCATCCGGAAGGAGGTCTCCACC
 TGGAGCCCTACATTGCTCTCTACCATGACTTCGTCACTGACTCAGAGGCTCAGAAAATTAGAGAACTTGCAGAAC
 CATGGCTACAGAGGTCACTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAGTACCGCATCAGCAAAAGTGCCT
 GGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCTCAACCACCGCATTGCTGCCCTCACAGGCCCTTGATGTCC
 GGCCTCCCTATGCAGAGTATCTGCAGGTGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATG
 CTACGTACCAAGCAGCCCCCTCTACAGAAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCT
 CGGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGAATGCAGCACTGT
 TTTGGTGGAACTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCATGCTGGCTGTCTGTCTCTGGTGGGAG
 ATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGGACAGGAATTCCGCAGACCCTGCAGCTCCAGCCCTGAAG
 ACT**GAA**CTGTTGGCAGAGAGAAGCTGGTGGAGTCTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTA
 GGAGAGGAGAAAGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCACTTTGTCTGTGCCTCGCAAATCAGAGGCAA
 GGGAGAGGTTGTTACAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAAGTCAGAGTAGGATGC
 ACAGTACAAAGGAGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGGAGTTCAGATACTCTCTGTTGGGAACAGG
 ACATCTCAACAGTCTCAGGTTTCATCAGTGGGTCTTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGG
 CAATGAGGACACCTGCAGGAGGGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGC
 AGCCCAAGCAGGGAGTGTCCCCCTCCCAGAAGCATATCCCAGATGAGTGGTACATTATATAAGGATTTTTTTTAA
 GTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACAAGTCATTAAAAATGTTTATAAATCAAAA

275/615

FIGURE 274

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64849
><subunit 1 of 1, 544 aa, 1 stop
><MW: 61126, pI: 6.40, NX(S/T): 2
MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLLGLLRRYLRGEEA
RLRDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGY
EKVEQDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRTVGSAITDLYSPKRLFSLT
GDDCFQVGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRA
GNVSCALSLSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNIPLQTRDT
YEGLCQTLGSQPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEA
QKIRELAEPWLQRSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRP
PYAEYLQVVNYGIGGHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLLSSVEAGGATAFIY
ANLSVPVVRNAALFWWNLHRSGEGSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSS
SPED

Important features of the protein:

Signal peptide:

Amino acids 1-19

Leucine zipper pattern:

Amino acids 34-56;41-63

Ribonucleotide reductase small subunit signature:

Amino acids 340-356

N-glycosylation sites:

Amino acids 242-246;482-486

Cell attachment sequence:

Amino acids 27-30

Tyrosine kinase phosphorylation site:

Amino acids 189-198

N-myristoylation sites:

Amino acids 4-10;135-141;153-159;164-170;241-247;303-309;309-315;
457-463;473-479

275/615 "275/615"

276/615

FIGURE 275

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCATCCTGGTGATCACC
TTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAAGCCAGGAAGCACAGCAAACGTCGAGTGAGAGAC
AAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTCTGGACAGAAGTCAATGCCTTGAAGGAAATTCAGCCCTG
CAGACAGTCTGTCTCCGAGGCACATAAGTTTCAACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCAT
GAGGCCAATGAAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACCTCCGACGAAATCAACGCCCTC
CAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGACATGGTCACGGAAGGC
AAGTTTGTGTGACGTCAACGGAATCGCTATCTCCTTCCCTCAACTGGGACCGTGCACAGCCTAACGGTGGCAAGCGA
GAAAACGTGTGTCTCTTCTCCCAATCAGCTCAGGGCAAGTGGAGTGATGAGGCCTGTGCGCAGCAGCAAGAGATAC
ATATGCGAGTTTACCATCCCTAAATAGGTCTTTCTCCAATGTGTCTTCCAAGCAAGATTTCATCATAACTTATAGG
TTCATGATCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAATGTCCAT
AGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGGCCCTTCTGGGGTATAGGGGATC
AGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCTAAACAGACTAAAATCTTTCTCTCTAGTCTTTC
TCACTTGTACAAACCCAGTTTGTCTTCAAAAAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAAGCAAGCTT
AGGCTACCTGAAAGATTTTCCCTTGGAAAGTTTAGCGTATGTTTGACTAACAAAAAATTCCTTACATCAGAGACTCT
AGGTGCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACCTTGTCTAGCC
CATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCTGCCATATCAGAACACAAACCCC
TGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCCTACCCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGA
AATGAAATGACAAGGTGTATATTTGATCAATTTTCATTCCCACCATTCGATTACAACCTCTAACTTAAATGGGTA
ACCTTAAGGCATATCAAAGAAGCAGATTGCATGATAAACGGAATAGAAAAAAGAACCTACATTTATTTTGCTT
TAGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCCTTTTTTACATTTTCGTATATTTAT
TTTTTTTTAGCCATCATTATATGTTTAAAGTCTATTATGGGCAACCAATCTTTGGAAGCTGAAAACCTGAATTTAAAG
AATGCTATCTTGGAAAATTGCATACGTCTGTGCAATTTTTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAG
ATTGTACAAAATAACTTCATTGCTTAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAG
CAAACAATTTTAAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATATCC
CACTTTGCAAACTTTAACTACACATGCTTGAATTAAGTTTTAGCTGTTTTCATTGCTCAATAATAAAGCCTGAA
TTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

20510-985501

277/615

FIGURE 276

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGLKTQIEKLWTEVNALKEIQALQT
VCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQDYGKRSLPGVNDFWLGINDMVTEGKF
VDVNGIAISFLNWDRAQPNGGKRENCVLFSSAQGKWSDEACRSSKRYICEFTIPK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GAGATAGGGAGTCTGGGTTTAAAGTTCTGCTCCATCTCAGGAGCCCCTGCTCCCAACCCCTAGGAAGCCACCAGAC
TCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCCACGAGCGCTGGCTGAGGGGACCGAGCC
GGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAGCGCCAGG**ATG**CCGCGCGGGGACTCGGAGCAGGTGCGCT
ACTGCGCGCGCTTCTCTCTACCTCTGGCTCAAGTTTTCACTTATCATCTATTCCACCCTGTTTCTGGCTGATTGGGG
CCCTGGTCTGTCTGTGCGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAACCCTTGAAAGTGCCTTCTCTGT
CTCCAGCCATCATCTCTCATCTCTCTGGGCGTGTGTCATGTTTCATGGTCTCTCTTTCATTGGTGTGCTGGCGTCCCTCC
GTGACAACTGTACCTTCTCCAAGCATTATGTACATCTTTGGGATCTGCCTCATCATGGAGCTCATTTGGTGGCG
TGGTGGCCTTGACCTTCCGGAACCAGACCATTGACTTCTGAAACGACAAACATTGAAGAGGAATTGAGAACTACT
ATGATGATCTGGACTTCAAAAACATCATGGACTTTGTTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACC
GAGATTTGGAGCAAGAATCAGTACCACGACTGCGAGTGCCCCCTGGACCCCCTGCCCTGTGGGGTGCCCTACACCTGCT
GCATCAGGAACAACAGCAGAAGTTGTCAACACCATTGTGTGGCTACAAACATATCGACAAGGAGCGTTTTCAGTGTGC
AGGATGTTCATCTACGTGCGGGGCTGCACCAACGCGGTGATCATCTGTGGTTTATGGACAACATACCCATCATGGCGT
GCATCCTCTTGGGCATCTGCTTCCCCAGTTTCTGGGGGTGCTGCTGACGCTGCTGTACATACCCCGGGTGGAGG
ACATCATCATGGAGCACTCTGTCACTGATGGGCTCTGGGGCCCCGTGCCAAGCCAGCGTGGAGGCGGCAGGCA
CGGGATGCTGCTTGTGCTACCCCAAT**TAG**GGGCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTCTGGGATAGC
ACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCTCTGCCCCACACTCAGTACTGACCAAGCCAGG
CTGTGTGTGGCTGTGTGTAGATCCCCAGCCCTCTGCCCTCCCCAGGAGCAGAGCTGGGCTCCCCTAAGAGG
TTTCCCGAGGCGAGTCTTGGAACTGTGCCCCACCTGGGGCTGGGGAAACAGGCCCTCTTTCTCCAGGCGCTGGG
CTACAGGGGAGGGAGAGCCTGAGGCTCTGCTCAGGGCCCATTTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCA
AGGCAGTTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCTCGGGGCAGGAGGGAAGGGCATCTGGGGAA
GGGCAGGAGGGAAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTGTGGCTCTTCTCAGCCTCCCAGGTGCCTT
GAGCCCTCTTGCAAGAGGGCGGCTGCTTCTTTCGAGCCTAGTTTTTTTTTACGTGATTTTTTGTAAACATTCATTTTTTT
GTACAGATAACAAGAGTTTCTGACTAATAAAGCTGGATTTTCCCCCATGTCTTATTCTTGCCCTTCCCCCAAC
CAGTTTTGTTAATCAAAACATAAAAAACATGTTTTGTTTTGTTTTTAAAAAAA

279/615

FIGURE 278

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863
><subunit 1 of 1, 294 aa, 1 stop
><MW: 33211, pI: 5.35, NX(S/T): 3
MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLESAFLAPAILILLGVVMFM
VSFIGVLASLRDNLVLLQAFMYILGICLIMELIGGVVALTFRNQTIDFLNDNIRRGIEYDDLDLDFKNIMDFVQK
KFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPTCCIRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAVII
WFMDNYTIMACILLGILLPQFLGVLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCLCYPN
```

Signal peptide:
amino acids 1-44

Transmembrane domains:
amino acids 22-42, 57-85, 93-116, 230-257

279/615

280/615

FIGURE 279

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACACCTGGGAAGAT
GGCCGGCCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATCCAAGCCACCCTCAGTCCCAC
TGCAGTTCTCATCCTCGGCCCCAAAAGTCATCAAAGAAAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAG
CATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCT
GGTGAACACCGTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCC
CTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAA
GACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGAAGTGGCCCCAC
CCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTATAAGCTCTCCTTCCT
GGTGAACGCCTTAGCTAAGCAGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTG
TCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCCATTTCCCTCAG
CATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTACAGCTCTACCTGGGGGCCAA
GTTGTTGGACTCACAGGGAAGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGA
CAACATCCCGTTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTATGGTCCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGA
AAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTTTTTTATAGA
CCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTTCCCTCCAGTGAAGCCCCTCCGCCCTTTGTT
CACCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAA
CATCAGCTCTGATCGGATCCAGCTGATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCAT
CACTGAGATCATCCACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCAATTGGT
GAAGGCCTTGGGATTGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCCTTGTG
GAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGGGTCCCAGCTGGGAGT
ATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAAACACTTGCCTGTGAAAA

281/615

FIGURE 280

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAMREKPAGGIPVLGS
LVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTI VEFHMTTEAQATIRMDTSASGP
TRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISL
SIDRLEFDLLYPAIKGDITQLYLGAALLDSQGKVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKA AVAAVLSPE
EFMVLLDSVLPESAHLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPL
FTLGIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSL
VKALGFEEAESLTKDALVLTPASLWKPPSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

281/615

282/615

FIGURE 281

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTCCCCGCGTTCT
CTTTCCACCTTTTCTTCTTCTTCCACCTTAGACCTCCCTTCCCTGCCCTCCTTTCCCTGCCACCGCTGCTTCCTGGC
CCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGGGTTGATCTGTGGCCCTGTGCCCTCCGTGTC
CTTTTCGTCTCCCTTCCCTCCCGACTCCGCTCCCGGACCAGCGGCTGACCTGGGGAAAAGGATGGTTTCCCGAGGT
GAGGGTCTCTCCTCCTTGTCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCGAGA
CATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCT
GATGTACTGCCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCTCCACTGTCCGCCTGTCCA
CTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCG
GGCCCCACCAAAGTCTTGCCAGCACAAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTT
CCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCACAACTG
CCCCGAACCAGGCTGCCCAGCACCCCTCCCCTGCCAGACTCCTGTGCCAAGCCTGCAAAGATGAGGCAAGTGA
GCAATCGGATGAAGAGGACAGTGTGCAGTGCCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGC
TGGGAGAAAAGAGAGGCCCCGGGCACCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGG
CGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTGCATCCTATG
CACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGTACCCCTGCCGTACCCCCGAGAA
AGTGGCTGGGAAGTGTGCAAGATTTGCCCAGAGGACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAG
GTGTCCCAAGGCACCGGGCCGGTCTCTGTCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGC
CCTGGAACACGAGGCCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCA
GAGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAGGAAGC
AAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCACTGGAACGTCTTCCTAG
CCCAGACCTTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGACCTAAACAGTTGCA
GATATGAGCTGTATAATTGTTGTTATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAA
AAAAAA

282/615

283/615

FIGURE 282

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902

><subunit 1 of 1, 451 aa, 1 stop

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHYPYLEPQGLMYCLRCTCSEGAHVSCYRLH
CPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELFP SRLPNQC VLCSCTEGQIYC
GLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVRHPQDPCSSDAGRKRGP GTPAPTGLSAPLSF
IPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPC
RHPEKVAGKCKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDE
ETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

283/615

CGCATGGTTCGCCCCGGTGGCGGTGGCGGCGGCGGTTGCGGAGGCTTCTTGGTTCGATTGCAACGAGGAGAAGAT
GACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGCCATGAGGAGCCTGCCGAGCCTGGGCG
GCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCGTCGCCTCAGCCGCCCTCGGCGGGGAATGTCACCGGTG
CGCGCGGGGCGCGGGCAGGTGAGCGCTCGCGGGGCCCGGGTTGCGGGCGGAGCCGACCCCTTCCCTA
GGGCGACGGCTCCCACGGCCAGGCCGAGGACCGGGCCCCCGCGCCGACCCGTCCACCGACCCCTGGCTGCGA
CTTCTCCAGCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCTCTTCCACCACCTTTCAGGCG
CGCTCGGCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACCTTCGACCACCTCTCAGGCGCCGACCAGACCCG
CGCCGACCACCCCTTTTCGACGACCACTGGCCCGGCGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCACGA
CTCCCGGAGACCCCGACCCCGATCTCCCAGCAGCAGCAACAGCAGCGCTCTCCCAACCCACCTGCCACCGAGG
CCCCCTCTTCGCCTCTCCAGAGTATGTATGTAAGTCTGTGGTTGGAAGCCTGAATGTGAATCGCTGCAACC
AGACCACAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCACTGTGAAACCTGCAAAGAGGGCTTTTACC
TAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCCACATGGAGCTCTCAGCATACCGTGCACAGGT
AAGCAACAGAGGGTGGAACGAAGTTTATTTTATTTTAGCAAGGGAAGGCTGCTACTCTCAAGGACCA
TACTGGTTTAAACAAAGGAGGATGAGGGTCATAGATTTACAAAATATTTATATACTTTATTTACTTTTAT
ATGTTATATTTAATGTGAGGATTTAAACATCTAATTTACTGATTTAGTTCTTCAAAGCACTAGAGTCGCCAA
TTTTTCTCTGGGATAATTTCTGTAAATTTATGGGAAAAATATTGAAGAATAAAATCTGCTTTCTGGAGGGCT
TTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTTATGTTTATTAATATACCATTGGAGTTTGAGGAAAT
TTGTTGTTTGGTTTATTTTTCTCTTAATCAAATTTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGG
GGTACCCTTAATTTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTGTGAGACCAAAA
GTAGATTAAAGCAGGAATTAATTTAACTATTATGTTATTTGGAGGTAATTAATCTAGTGAATAATGTACTGT
TATCTAAGCATTGGCTTGTACTGCACTGAAAGTAATTAATCTTTGACCTTATGTGAGGCATCTGGCTTTTGTG
GACCCCAAGTCAAAAAAAGTGAAGAGACAGTATTAATAATGAAAAAATAATGACAGGTTATACTCAGTGTAACC
TGGGTATAACCCAAGATCTGCTGCCACTTACGAGCTGTGTTCTTTGGGCAAGTAATTTCCCTTTCAGTGAAGCT
TTCTTCTCAAGGTTGTTGTGAAGATTAATGAGTTGATATATATAAAATGCCCTAGCACATGTCACCTCAATAAAAT
CTGGTTTGTGTTTAAATTTCAAAGGAATATTATGCACTGAAATGAGAGAACATGTTTAAAGAACTTTTAGCTCTTTG
ACAAAGAAGTGCCTTTATACCTTTAGCACTAAATATTTTAAATGCTTTATAAAATGATATTATACTGTTATGGAATAT
TGTATCATATTGTAGTTTATTAATAATGTAGAAGAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTT
GGGAGGCCAAGGCGGGTGGATCACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCCGTCT
CTACTAAAAATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATCGGTTGAACCCGGGAGGTGGAGTTGCAGTGAGCTGAGATCGCGCCACTGCACCTCAGCCTGGTGAG
AGAGGGAGACTCTGCTCTTAAAAAAGGAAAAAAGGAAAAA

285/615

FIGURE 284

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952
><subunit 1 of 1, 258 aa, 1 stop
><MW: 25716, pI: 8.13, NX(S/T): 5
MRSLPSLGGLALLCCAAAAA AVASAASAGNVTGGGGAAGQVDASPGPGLRGEPSHPFFPRATAPTAQAPRTGPPRA
TVHRPLAATSPAQSPETTTPLWATAGPSSSTTFQAPLGPSTTPPAAERTSTTSQAPTRPAPTTLSTTTGPAPTPV
ATTVPAPTTTPRTPTPDLPSSSNSSVLPTPPATEAPSSPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHC
ETCKEGFYLNYSGLCQPCDCSPHGALSIPCNR
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AACAGACGTTTCCCTCGCGGCCCTGGCACCTCTAACCCACAGACATGCTGCTGCTGCTGCTGCCCTGCTCTGGGG
AGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCGGTGACGGTGCAGGAAGGCCTGTGT
GTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATTACCCCTGGCCAGTAGTTTCATGGCTACTGG
TTCCGGGAAGGGGCCAATACAGACACGGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAG
ACTCTGGGACCGGATTCACCTCCTTGGGGAGCCACATACCAAGAATTGCACCTTGCAGCATCAGAGTAGCCAGAAGA
AGTGATGCGGGGAGATACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCT
GTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCCCCCAGAAT
CTGACCTGCTCTGTGCCCTGGGCCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATAGGGACCTCCGTGTCC
CCCTGGACCCCTCACCACCCGCTCCTCGGTGCTCACCCCTCATCCCACAGCCCCAGGACCATGGCACCCAGCCCTC
ACCTGTCCAGGTGACCTTCCCTGGGGCCAGCGTGACACGAACAAGACCGTCCATCTCAACGTGTCTTACCCCGCTC
CAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCAGTAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCA
TCTCCACAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCCTGAGC
CTGAGCTGGAGAGGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGCTCTCC
CTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTTCGGGGGAGCTGGAGCCACAGCCCTGGTCTTCTGT
TCTTCTGCGTCACTCTTCTGTAGTGAGGTCTTCAGGAAGAAATCGGCAAGGCCACGAGCGGGCTGGGGAGAT
ACGGGCATAGAGGATGCAAACGCTGTTCAGGGGTTGAGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGAC
AGTCCCCCAGACCAGCCTCCCCCAGCTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTC
AGCTTCCAGATGGTGAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATC
CACAGATGAAGAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCCCTCAGGCAAGGGAGAAGTCAGAGGC
TGATTTCTGTAGAAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTTATGTGCAGAGTGAAGAAG
ACACAGGCTTTAGAGTCAAAGTATCTCAAAGCTGAATCCACACTGTGCCCTCCCTTTTATTTTTTTAACTAAAG
ACAGACAAATTCCTA

287/615

FIGURE 286

MLLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREGANTDQDAPVA
TNNPARAVWEETRDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNYKHHRLSVNVTALTHRPNILI
PGTLESGCPQNLTCVSPWACEQGTPPMISWIGTSVSPLDPSTTRSSVLTLLIPQPDHGTSLTCQVTFFPGASVTTN
KTVHLNVSYPPQNLMTVFQGDGTVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPS
NPGVLELPWVHLRDAAEFTCRAQNPLGSQQVYLNVSLSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRK
KSARPAAGVGDTGIEDANAVRGASASQGPLEPWAEDSPPDQPPASARSSVGEDELQYASLSFQMKPWDSRGQE
ATDTEYSEIKIHR

Signal peptide:
amino acids 1-15

Transmembrane domain:
amino acids 351-370

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 287

CGCGAGCTGAGAGGAGCAGGTAGAGGGGCGAGGGCGGGACTGTCGTCTGGGGGAGCCGCCAGGAGGCTCCTCAG
 GCCGACCCAGACCCCTGGCTGGCCAGGATGAAGTATCTCCGGCACCGGCGGCCCAATGCCACCCCTCATTTCTGGCC
 ATCGGCGCTTTTACCCTCCTCCTCTTCACTGTCTAGTGTACACCCACCTGCAAGGTCCAGGAGCAGCCACCG
 GCGATCCCCGAGGCCCTGGCTGGCCCACTCCACCCACCCGCCCAGCCCCGGCCCCGTGCCATGCCAACACCTCT
 ATGGTCACCCACCCGGACTTCGCCACGCAGCCGAGCAGCTTCAGAACTTCCTCCTGTACAGACACTGCCGCCAC
 TTTCCCTGCTGTCAGGACGTGCCCCCTCTAAGTGCGCGCAGCCGGTCTTCCTGCTGCTGGTGATCAAGTCTCTCC
 CCTAGCAACTATGTGCGCCGCGAGCTGCTGCGGCGCACGTGGGGCCGCGAGCGCAAGGTACGGGGTTTGCAGCTG
 CGCTCCTCTTCCTGGTGGGCACAGCCTCCAACCCGACAGAGGCCGCAAGGTCAACCGGCTGCTGGAGCTGGAG
 GCACAGACTCACGGAGACATCCTGCAGTGGGACTTCCACGACTCCTTCTTCAACCTCACGCTCAAGCAGGTCTCTG
 TTCTTACAGTGGCAGGAGACAAGGTGCGCCAACGCCAGCTTCGTGCTCAACGGGGATGATGACGTCTTTGCACAC
 ACAGACAACATGGTCTTCTACCTGCAGGACCATGACCCTGGCCGCCACCTCTTCGTGGGGCAACTGATCCAAAAC
 GTGGGCCCCATCCGGGCTTTTGGAGCAAGTACTATGTGCCAGAGGTGGTGACTCAGAATGAGCGGTACCCACCC
 TATTGTGGGGGTGGTGGCTTCTTGCTGTCCCGCTTCACGGCCGCTGCCCTGCGCCGTGCTGCCCATGTCTTGGAC
 ATCTTCCCCATTGATGATGTCTTCTGGGTATGTGTCTGGAGCTTGAGGGACTGAAGCCTGCCTCCCACAGCGGC
 ATCCGCACGTCTGGCGTGCGGGCTCCATCGCAACACCTGTCTCTTTGACCCCTGCTTCTACCGAGACCTGCTG
 CTGGTGCACCGCTTCTTACCTTATGAGATGCTGCTCATGTGGGATGCGCTGAACCAGCCCAACCTCACCTGCGGC
 AATCAGACACAGATCTACTGAGTCAGCATCAGGGTCCCCAGCCTCTGGGCTCCTGTTTCCATAGGAAGGGGCGAC
 ACCTTCTCTCCAGGAAGCTGAGACCTTTGTGGTCTGAGCATAAGGGAGTGCCAGGGAAGGTTTGAGGTTTGATGA
 GTGAATATTCTGGCTGGCGAACTCCTACACATCCTTCAAAACCCACCTGGTACTGTTCCAGCATCTTCCCTGGAT
 GGCTGGAGGAACCTCAGAAAATATCCATCTTCTTTTGTGGCTGCTAATGGCAGAAGTGCCTGTGCTAGAGTTCC
 AACTGTGGATGCATCCGTCCCGTTTGAGTCAAAGTCTTACTTCCCTGCTCTCACCTACTCACAGACGGGATGCTA
 AGCAGTGCACCTGCAGTGGTTTTAATGGCAGATAAGCTCCGTCTGCAGTTCCAGGCCAGCCAGAACTCCTGTGTC
 CACATAGAGCTGACGTGAGAAATATCTTTCAGCCCAGGAGAGAGGGGTCTGATCTTAACCCCTTCTCTGGGTCTC
 AGACAACTCAGAAGGTTGGGGGGATACAGAGAGGTGGTGAATAGGACCGCCCCCTCTTACTTGTGGGATCAA
 ATGCTGTAATGGTGGAGGTGTGGGCAGAGGAGGGAGGCAAGTGTCTTTGAAAGTTGTGAGAGCTCAGAGTTTCTG
 GGGTCTCTCATTAGGAGCCCCCATCCCTGTGTTCCCCAAGAATTACAGAGAACAGCACTGGGGCTGGAATGATCTTT
 AATGGGCCCCAAGGCCAACAGGCATATGCCTCACTACTGCCTGGAGAAGGGAGAGATTGAGGTCTCTCAGCAGCCT
 CCTCACCCAGTATGTTTTACAGATTACGGGGGACCGGGTGAGCCAGTGACCCCTGCAGCCCCCAGCTTCAGG
 CCTCAGTGTCTGCCAGTCAAGCTTCACAGGCATTGTGATGGGGCAGCCTTGGGGGAATATAAAATTTTGTGAAGAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

289/615

FIGURE 288

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65413

<subunit 1 of 1, 372 aa, 1 stop

<MW: 42515, pI: 8.92, NX(S/T): 6

MKYLRHRRPNATLILAIGAFTLLLFSLLVSPPTCKVQEQPPAIPALAWPTPPTRPAPAPCHANTSMVTHPDFAT
QPQHVQNFLLYRHCRHFPLLQDVPPSKCAQPVFLLLVIKSSPSNYVRRELLRRTWGRERKVRGLQLRLLFLVGTA
SNPHEARKVNRLLELEAQTHGDILQWDFHDSFFNLTLKQVLFQWQETRCANASFVLNGDDDVFAHTDNMVFYLO
DHDPRGRLHFLVQGQLIQNVGPIRAFWSKYVPEVVTQNERYPYPCGGGGFLLSRFTAAALRRAHVLDIFPIDDVFL
GMCLELEGLKPASHSGIRTSGVRAPSQHLSSFDPFCFYRDLLLVRFLPYEMLLMWDALNQPNLTCGNQTQIY

Important features:

Type II transmembrane domain:

Amino acids 15-34

N-glycosylation sites:

Amino acids 10-14;64-68;184-188;202-206;362-366;367-371

TonB-dependent receptor proteins signature 1:

Amino acids 1-32

N-myristoylation sites:

Amino acids 308-314;316-322

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 289A

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGAGAGGCAGCCTCCTCCAGGAGCGGGGCCCT
GCACACC**ATG**CCCCCGGGTGGGCAGGGGTTCGGCGCCGCGCGTGCAGCGCCGCGCTGGCGCTGGCCTTGCGCTGGC
GAGCGTCCCTGAGTGGGCCTCCAGCCGTGCGCTGCCCCACCAAGTGTAACCTGCTCCGCTGCCAGCGTGGACTGCCA
CGGGCTGGGCCTCCGCGCGGTTCCTCGGGGCATCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATAT
CACCAGGATCACCAAGATGGACTTCGCTGGGCTCAAGAACCTCCGAGTCTTGCACTCTGGAAGACAACCAGGTCAG
CGTCATCGAGAGAGGCGCCTTCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACAAGAATAAGCTGCAAGT
CCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTAGATTTGAGTGAAAACCAGATCCAGGGGAT
CCCCGAGGAAGGCGTTCCGCGGCATCACCGATGTGAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGA
AGATGGAGCCTTCCGAGCGCTGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGT
CACCAGCTTCAACCACATGCCGAAGATCCGAACCTGCGCCTCCACTCCAACCACCTCTACTGCGACTGCCACCT
GGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTCACACTCTGCATGGCTCCTGTGCATTT
GAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTGTGCCAGCCCCCCTCGGAGCCCCCATCCTG
CAATGCCAACTCCATCTCCTGCCCTTCGCCCCGACGTGCAGCAATAACATCGTGGACTGTGAGGAAAGGGCTT
GATGGAGATTCTGCCAACTTGCCGGAGGGCATCGTCGAAATACGCCCTAGAACAAGAACTCCATCAAAGCCATCCC
TGCAGGAGCCTTCAACCAGTACAAGAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGATATTGCTCC
AGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCTGTATGGGAACAAGATCACCGAGATTGCCAAGGG
ACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTCAATGCCAACAAGATCAACTGCCTGCGGGTGAACAC
GTTTCAGGACCTGCAGAACCTCAACTTGCTCTCCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGCTCTT
CGCCCCCTGTCAGTCCATCCAGACACTCCACTTAGCCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCT
GGCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCGCTGCAGCAGCCCCGCGCGACTCGCCAACAA
GCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGGATTACCGCAGCAGGTTTCAGCAGCGA
GTGCTTCATGGACCTCGTGTGCCCCGAGAAGTGTGCTGTGAGGGCACGATTGTGGACTGCTCCAACCAGAAGCT
GGTCCGCATCCCAAGCCACCTCCCTGAATATGTACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGA
GGCCACTGGCATCTTCAAGAAGTTGCCCAACCTGCGGAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTGCG
AGAGGGAGCTTTTCGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCTGGAGACCGTGCACGG
GCGCGTGTTCGTGGCCTCAGTGGCCTCAAAACCTTGATGCTGAGGAGTAACCTTGATCAGCTGTGTGAGTAATGA
CACCTTTGCCGGCCTGAGTTCCGGTGAAGTGTGCTGCTCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGC
CTTACCACGCTTGTCTCCCTGTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTG
GCTCGGCAAGTGGTTGAGGAAGAGGCGGATCGTCAGTGGGAACCCCTAGGTGCCAGAAGCCATTTTCTCAAGGA
GATTTCCATCCAGGATGTGGCCATCCAGGACTTCACCTGTGATGGCAACGAGGAGAGTAGCTGCCAGCTGAGCCC
GCGCTGCCCCGAGCAGTGCACCTGTATGGAGACAGTGGTGGGATGCAGCAACAAGGGGCTCCGCGCCCTCCCCAG
AGGCATGCCCCAAGGATGTGACCGAGCTGTACCTGGAAGGAAACCACCTAACAGCCGTGCCAGAGAGCTGTCCGC
CCTCCGACACCTGACGCTTATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCAGTAACAT
GTCTCACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCTTCAACGGGCTGCG
GTCCCTGCGAGTGCTAACCTCCATGGCAATGACATTTCCAGCGTTCCTGAAGGCTCCTTCAACGACCTCACATC
TCTTTCCATCTGGCGCTGGGAACCAACCACTCCACTGTGACTGCAGTCTTCGGTGGCTGTGCGAGTGGGTGAA
GGCGGGGTACAAGGAGCCTGGCATCGCCGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCAC
CCCAACCCACCGCTTCAGTGCAAAGGGCCAGTGGACATCAACATTGTGGCCAAATGCAATGCCTGCCTCTCCAG
CCCGTGCAAGAATAACGGGACATGCACCCAGGACCTGTGGAGCTGTACCGCTGTGCCTGCCCC

TACAGCTACAAGGGCAAGGACTGCACCTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCAGCATGGAGGCACC
TGCCACCTGAGTGACAGCCACAAGGATGGGTTCAGCTGCTCCTGCCCTCTGGGCTTTGAGGGGCAGCGGTGTGAG
ATCAACCCAGATGACTGTGAGGACAACGACTGCGAAAAAATGCCACCTGCGTGAGACGGGATCAACAACTACGTG
TGTATCTGTCCGCCTAACTACACAGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTC
TGTGAGCATGAGGCCAAGTGCATCCCCCTGGACAAGGATTCAGTGCAGTGTGTCTCTGGTTCACAGCGGGAAC
CTCTGTGAGACAGACAATGATGACTGTGTGGCCCAAGTGCAGGACGGGGCCAGTGCCTGGACACAATCAAT
GGCTACACATGCACCTGCCCCCAGGGCTTCAGTGGACCCCTCTGTGAACACCCCCACCCATGGTCTTACTGCAG
ACCAGCCCATGCGACCAGTACGAGTGCCAGAACGGGGCCCAGTGCATCGTGGTGACAGCAGGAGCCACCTGCCGC
TGCCACACAGGCTTCGCCGGCCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAGAGCTCCTACGTG
GAACTGGCTCTCCGCAAGGTCCGACCCCCAGGCCAACATCTCCCTGCGAGGTGGCCACTGACAAGGACAACCGGCATC
CTTCTCTACAAGGACACAATGACCCCCCTGGCAGCTGACAGGCGCCAGTGCAGCTGGTCTATGACAGC
CTGAGTTCCTCCCTCCAACCAAGTGTACAGTGTGGAGACAGTGAATGATGGGCAGTTTTCAGAGTGTGGAGCTGGTG
ACGCTAAACCCAGACCTGAACTAGTAGTGGACAAAAGAACTCCAAAGAGCCTGGGGGAAGCTCCAGAAGCAGCCA
GCAGTGGGCATCAACAGCCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCTCTCCGCTTGCGCCAGGGC
ACGGACCGGCCTCTAGGCGGCTTCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAGGACTTCAAG
GCCCTCCCACCACAGTCCCTGGGGGTGTCAACAGGCTGCAAGTCTGACCGGTGTGCAAGCACGGCCTGTGCCGC
TCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGGCCGCCAGGCTGGACCGGCCACTCTGCGACCAAGGAGGCCCG
GACCCCTGCCTCGGCCACAGATGCCACCATGAAAAATGTGTGGCAACTGGGACCTCATACTGTGCAAGTGTGCC
GAGGGCTATGGAGGGGACTGTGTGACAAACAAGAAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTACCAT
GGGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGCGAGCACTGCCAA
CAAGAGAATCCGTGCCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCAGAAAGTTATGCATCATGTGCCACA
GCCTCCAAGGTGCCCATCATGGAATGTGCTGGGGGCTGTGGGGCCCAGTGCTGCCAGCCCACCCGCAGCAAGCGG
CGGAAATACGTCTTCCAGTGACCGGACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGC
CTCGCGTGTTCTTCAAGGCCCTGCCCGCTGCCTGCCACCTCTCGGACTCAGCTTATGATGGAGTTGGGACAGCCAT
GTGGGATCCCCCTGGTGTCTCAGCATGAAGGAATGAAGCTGGAGAGGAAGCTAAAGAAGAAGAGAATATTAACTA
TATTGTAAAAATAAACAAAAAATAGAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

292/615

FIGURE 290

MAPGWAGVGAAVRARLALALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRNAERLDLDRNNITR
ITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRNLNKNKLQVLPELLFQSTPKLTRLDLSENQIQGIPR
KAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEILTNNNNNISRIILVTSFNHMPKIRTLRLHSNHLYCDCHLAW
LSDWLRQRRRTVGQFTLCMAPVHLRGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLME
IPANLPEGIVEIRLEQNSIKAIPAGAFTQYKKLKRIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIAKGLF
DGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTIKGLFAPLQSIQTLHLAQNPVFCDCHLKWLAD
YLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDYRSRFSSECFMDLVCPEKCRCEGTIVDCSNQKLVR
IPSHLPEYVTDLRRLNDNEVSVLEATGIFKKLPNLRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRV
FRGLSGLKTLMLRSNLISCVSNDTFAGLSSVRLLSLYDNRIITITPGAFTTLVSLSTINLLSNPFNCNCHLAWLG
KWLKRRIIVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSQLSPRCPEQCTCMETVVRC SNKGLRALPRGM
PKDVTLEYLEGNHILTAVPRELSALRHLTLIDLSNNSISMLTNYTFSNM SHLSTLILSYNRLRCIPVHAFNGLRSL
RVLTTLHGNDISSVPEGSFNDLTSLSHLALGTNPLHCDCLRWLSEWVKAGYKEPGIARCSSPEPMADRLLLLTTPT
HRFQCKGPVDINIVAKCNACLSSPCKNNGTCTQDPVELYRCACPYSYKGKDCVPIINTCIQNPCQHGGTCHLSDS
HKDGFSCSCPLGFEGQRCIEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPENLNCQHEAK
CIPLDKGFSCCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFSGPFCEHPPPMVLLQTSPCDQ
YECQNGAQCIIVVQOEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRPQANISLQVATDKDNGILLYKGD
NDPLALELYQGHVRLVYDSLSPPTTVYSVETVNDGQFHSVELVTLNQTLNLVVDKGT PKSLGKLQKQPAVGINS
PLYLGGIPTSTGLSALRQGTDRPLGGFHGCIHEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDS
VVCECRPGWTGPLCDQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKND SANACSAFKCHHGQCHIS
DQGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQPTRSKRRKYVFQ
CTDGSSFVEEVERHLECGCLACS

Signal peptide:

amino acids 1-27

292/615

293/615

FIGURE 291

GGATGCAGGACGCTCCCCTGAGCTGCCTGTACCGACTAGGTGGAGCAGTGTTTCTTCCGCAGACTCAACTGAGA
AGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCTCCGGCAGGCTTTGAGG**ATGA**AGGCTG
CGGGCATTCTGACCCCTCATTTGGCTGCCTGGTACAGGCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAA
AAATATTCTCGAGGGCTGGCCTGGACAATTACTGGGGCTTCAGCCTTGGAACCTGGATCTGCATGGCATATTATG
AGAGCGGCTACAACACCACAGCCCCGACGGTCTTGATGACGGCAGCATCGACTATGGCATCTTCCAGATCAACA
GCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACCAACTGCCATGTGCGCTGCTCAGCCTTGATCACTG
ATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGAGACACAAGGAATGAACTATTGGCAAGGCT
GGAAGAAACATTTGTGAGGGCAGAGACCTGTCCGAGTGGAAGGCTGTGAGGTTTCC**TAA**ACTGGAACCTGGAC
CCAGGATGCTTTGCAGCAACGCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTATCTTGTCCCGTTTCCT
CCCAATATTCTTCTCAAACCTGGAGAGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCATTTAAATGTC

2025-09-01 10:00:00

294/615

FIGURE 292

MKAAGILTTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAPTVLDDGSIDYGIF
QINSAWCRRGKCLKENNHCHVACSALITDDLTDALICARKIVKETQGMNYWQGKKHCEGRDLSEWKKGCEVS

Signal peptide:
amino acids 1-19

294/615

295/615

FIGURE 293

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTACCCAAGGAAAAG
TGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTCTCATAGCGACCACCAGAGGA
TGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGC
AAGGAAATCAAAGACGAATGTCCTAGTGCAATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTAC
CAGACCTTCTGTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGT
GGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAAC
TGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACTACGAC
ATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCCAATAAGTCCCCCATGCAGCACTGGAGAAAACAGCTCCCTG
CTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCA
GTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGCGACGCC
CAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTT
AATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATT
GGTGGAGGAGGATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACCTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGTTGAGAGTTT
TGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCAGTAGCTAGA
ATGTTAATGGCAGAAGAGAAAAACAATAAATCATATTGACTCAAGAAAAAAA

295/615

296/615

FIGURE 294

MNQLSFLFLFIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVIYQTFCDMTSG
GGGWTILVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAATSDDYKNPGYYDIQAKDLGIW
HVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYISP
YGQREFTAGFVQFRVFNNERAANALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSS
REITEAAVLLFYR

2025-10-26 10:56:00

CAGGCCATTTGCATCCAC TGTCTTGTGTTCGGAGCCAGGCCACACCGTCTCAGCAGTGT CATGTGTTAAAAA
CGCCAAGCTGAATATATCATGCCCCCTATTAAAACTTGTACATGGCTCCCCATTGGTTTTTGGAGAAAAGTTCAG
CTTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTCAGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTTACTGT
CACTTCCCAGATCTGCTTCTCAACAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATA
GAGGAATACACCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTTCAGAGTGTGTTTTTAAAGGAACA
TTCAAGTCCCTGTCTCACATAGACCCCTGATGTCTCTATCCATCTCTAAATGTACCAGCTTTGACTCAGTTGTT
CCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAAATTCTGTGCTCTCCATCAACAGATACGAAAGGAAG
AAAAATCTGACTTTGGCAC'TGGAAGCCCTAGTACAGCTGCGTGGAGATTGACATCCCAAGATTGGGAGAGGGTT
CATCTGATCGTGGCAGGTGGTTATGACGAGAGAGTCTTGGAGAATGTGGAACAT'TACAGGAAT'TGAAGAAATG
GTCCAAACAGTCCGACCTTGGCCAGTATGTAGACTTTCTTGAGGCTTTCTCAGACAAACAGAAATCTCCCTCTC
CAGCACTGCACGTGTGTGCTTTTACACACCAAGCAATGAGCACTTGGCATTGTCCCTCTGGAAGCCATGTACATG
CAGTGCCCAAGTCATTGCTGTTAATTGGGTGGACCCCTTGGAGTCCATTGACCACAGTGTACAGGGTTTTCTGTGT
GAGCCTGACCCGGTGCATTTCTCAGAAGCAATAGAAAAGTTTATCCGTGAACCTTCTTTAAAGCCACCATGGGC
CTGGCTGGAAGAGCCAGAGTGAAGGAAAAAATTTTCCCTGAAGCATTTTACAGAACAGCTCTACCCGATATGTTACC
AACTGCTGGTATATATCAGATTGTTTTTAAGATCTCCATTAAATGTCATTTTATGGATTGTAGACCCAGTTTTGA
AACCAAAAAAGAAACCCTAGAACTTAATGCAAGAGATCTTTTAAAAAATAAACTTGAGTCTTGAATGTAGGCCA
CTTTCCTATATACCACACCTCCCTGTCCACTTTTTCAGAAAAAACCATGTCTTTTATGCTATAATCATTCCAAATTT
TGCCAGTGT'TAAGTTACAAATGTGGTGTCA'TTCCATGTTT CAGCAGAGTATTTTAATTATATTTTTCTCGGGATTAT
TGCTCTTCTGTCTATAAAATTTTGAATGATCTGTGCCTTAATTGGTTTTTCATAGTTTTAAGTGTGTATCATTATCA
AAGTTGATTAA'TTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTTCCAGATTCAATCCACCGAAGTGT
TCACTGTCACTGTGTTAGGGAATTTTTGTTTGTCTGTCTTGTGCTGGATCCATAGCGAGAGTGCTCTGTATTTTT
TTAAGATAA'TTTGTATTTTTTGCACTGAGATATAATAAAAGGTGTTTATCATAAAAAAAAAAAAAAAAAAAAAA

298/615

FIGURE 296

MPLLKLVHGSPLVFGEKFKLFTLVSAIPVFRLLARRRKILFYCHFPDLLLLTKRDSFLKRRLYRAPIDWIEEYTTG
MADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPEKLDDLVPKGKKFLLLSINRYERKKNLTLA
LEALVQLRGRLLTSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQYVTFLLRSFSDKQKISLLHSCTCV
LYTPSNEHFGIVPLEAMYMQCPVIAVNSGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRAR
VKEKFSPEAFTEQLYRYVTKLLV

Signal peptide:

amino acids 1-15

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 297

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACCATGTTGGACTTCGCGATCTTCGCCGTTACCTTC
TTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGACAAGCTGCAGGAATTCCAGGGATTACT
CCAACTGAAGAAAAAGATGGTAATCTTCCAGATATTGTGAATAGTGAAGTTTGCATGAGTTCCTGGTTAATTTG
CATGAGAGATATGGGCCTGTGGTCTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTA
CTGAAGCAGCATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTATCAA
TCTGGTGGTGGCAGTGTGAGTGAAGAACCATGAGGAAAAAATTGTATGAAATGGTGTGACTGATTCTCTGAAG
AGTAACTTTGGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATAAATGGCTCTCCTACCCAGAGACCCAGCAC
GTGCCCCCTCAGCCAGCATATGCTTGGTTTTGCTATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAA
GATGATCAGGAAGTCATTTCGCTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTTCTAGAT
GGGTCACTTGATAAAAAACATGACTCGGAAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTTTTAAGG
AACATCATAAAAGAACGAAAAGGAAGGAACCTTCAGTCAACATATTTTCATTGACTCCTTAGTACAAGGGAACCTT
AATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGGCCAGTTGCATAATAACTGCAAAATTGTGTACC
TGGGCAATCTGTTTTTTAACCACCTCTGAAGAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTGGA
AATGGTCCTGTTACTCCAGAGAAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTCGAACT
GCCAACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATTCCTAGAGAG
ACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCCATCTCCACACAAGTTTGATCCA
GATCGGTTTGTATGATGAATTAGTAATGAAAACCTTTTTCCTCACTTGGATTCTCAGGCACACAGGAGTGTCCAGAG
TTGAGGTTTGCATATATGGTGACCACAGTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAG
GGACAGGTTATTGAAACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGA
TATTAATAATTTTATACATTTAAATCATTTGTTAAATTGATTGAGGAAAAACAACCATTTAAAAAAAATCTATGTTG
AATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTACTTAA

300/615

FIGURE 298

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLHERYGPVVSFWFGR
RLVVS LGTV DVLKQHINPNKTS DPFETMLKSLRYQSGGGSVSENHMRKKLYENGVTDSLKSNFALLKLSEELL
DKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFEDDQEVIRFQKNHGTWVSEIGKGFLDGS LDKNMTRKKQYE
DALMQLESVLRNIIKERKGRNFSQHIFIDSLVQGNLNDQQILED SMIFSLASCIITAKLCTWAICFLT TSEEVQK
KLYEEINQVFGNGPVTPEKIEQLRYCQHVLCE TVRTAKLTPVSAQLQDIEGKIDRFIIPRET LVLYALGVVLQDP
NTWPSPHKFD PDRFDDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLSVEGQVIETKYELVTSS
REEAWITVSKRY

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 271-290

300/615

301/615

FIGURE 299

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGACCGCCGCCCTT
GTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACGCTCCTGGCGCATCTGGTGGT
CGTCATCACCTTATTCTGGTCCCAGGACAGCAACATACAGGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTA
TGACAAGCAGGACATTTCAGCTGGTGGCCGCGCTCTCTGTCACCTTGGGCCTCTTTGCAGTGGAGCTGGCCGTTT
CCTCTCAGGAGTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGC
CCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCC
AGCTGTCACCTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTTGAATTACCTTCATGACG
GGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAGGCATAGGCTTCGGTT
TTCCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATT
GTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACAATTAAA
AAAAAAAAA

20570 "9865001

302/615

FIGURE 300

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFAVELAGFLSGV
SMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

Downloaded from www.ascp.org

303/615

FIGURE 301

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGACTCGCTGCTGC
TTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATGGTCCTCCCGGCTCAGAGGAC
CCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCC
CGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCC
CCCAACCGCCCCGAACCACAGCCCCCACCCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCC
AACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTC
CACTTCCAACACAATGCCACAGGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAGAGTTC
CACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTA
GAACGGGGCCGCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCA
GCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTATCGGCTG
GTCCAGAAGGTGTGCCAGATTACAATAACCATAGTGATAACCCCTACTACCCATCTGGGTGAACCGGGGCAGGC
CACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGG
GTTGGGCCTCAGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTG
GGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCCTGTGGGCAG
GCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTTCTTGCCATCCTGAGGAAAGATA
GCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAG
GAGCCAGTCAGCAGGGTGGGGTGGGGCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTC
GTGTGCTGAGCATGGCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCT
CCAGCCAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATGGCACA
CCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCCACTAAGGCCACAGCCCATCCGCGTGCTGTGTGTC
CCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCC
TCACCTGTCAGACCGGGTTCTCCCGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGC
CGGGCCCGCAGAGCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGTGA
AACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

303/615

304/615

FIGURE 302

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

><subunit 1 of 1, 252 aa, 1 stop

><MW: 28127, pI: 8.91, NX(S/T): 5

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTLLGLLAPPGEA
WGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTFSVHFQHNATGQGNISISLVP
PSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDDPAKICSRDHAQSSATWSCSQPFKVVVCVYIAF
YSTDYRLVQKVCPCDYNHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

305/615

FIGURE 303

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCCCTTTATGTCT
TCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTTCTGGTTGGTGTCTCTACTGA
TTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACAAAGATGGACCAACACAGAAATATCTGCTGA
TCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCA
GTGAAGGTTTGAAGAGTATAAACCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGG
GCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCA
TTCATGGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGTATTCT
GGGGCATTGTATTTTGTATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGC
TGGTGTCTAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCA
TGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAACTCTGCCTGCTCTGCCAAGACAAGA
ACTTTCTTCTTTACAACCAGCGCTCCAGATAACCTCAGGGAACCAGCAC'TTCCCAAACCGCAGACTACATCTTTA
GAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

2025-10-28 10:55:00

306/615

FIGURE 304

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658

><subunit 1 of 1, 257 aa, 1 stop

><MW: 28472, pI: 9.33, NX(S/T): 0

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKGPTQKYLLIFG
AFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGVFSFVNTLSDSLGPSTVGIHG
DSPQFFLYSAFMTLVIILLHVFWGIVFFDGCCKKKGILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGT
WAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

307/615

FIGURE 305

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTGGACCCAAAGG
TAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAACTCGGTTCATTACCACAGCTCA
AACCTGCTTTGGGACTCCCTCCCAAAAAGTGGCTCCGGATCAGGGAACACTACCAAACCAACAGCAGTCAAATC
AGGTCTTTCCTTCTTTAAGTCTGATACCATTAAACACAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATC
CTGCTGCAGGAATGACACCTGGTACCCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGC
ACCCACATGTGTTACCAATTTTGTGCACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCAC
AAATCTTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGGCTAATC
CAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGGGAACCCCAGCAGGGCC
GCCTCCCAACTCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCTGCAGGCATCCAAAGGAGCACACATG
CCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCAGTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCT
CGAATTTGGTGATACATGTGAATCTTTATCATTGATTATATTTATGGAATAGATTGAGACACATTGGATAGTCTTA
GAAGAAATTAATTCTTAATTTACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAA
CTTTTAAAAACAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGAAAACTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

2025-10-10 10:54:50

308/615

FIGURE 306

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

310/615

FIGURE 308

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNRPVDVLVPSVSLQA
FKSFLRSQGLEAYVTIEDLQALLDNEDEMQHNEGQERSNNFNFGAYHSLEAIYHEMDNIAADFPDLARRVKIG
HSFENRPMYVLKFSTGKGVRPAPVWLNAGIHSREWISQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANP
DGYVYTQTQNLWRKTRSRNPGSSCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVDIFIQKHGN
FKGFIDLHSYSQLLMYPYGYSVKKAPDAEELDKVARLAALKALASVSGTEYQVGPTCTTVYPASGSSIDWAYDNGI
KFAFTFELRDTGTYGFLLPANQIIPTAETWLGLKTIMEHVVDNLY

Signal peptide:
amino acids 1-16

2025-09-09 10:00:00

311/615

FIGURE 309

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCACGCCTGAGTCCAAGAT
TCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTATCTCTTCACCTTCAAGTCCC
CTTTCTCAAGAATCCTCTGTTCCTTTGCCCTCTAAAGTCTTGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGC
CACAAAGAGACAGATGAAGATGCAGAAAGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGC
AACAAATTCCAATGAGACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCAC
CAACTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGG
GGTCAGCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAGTTCAG
CACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGG
AGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAG
CACACTCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCAC
CAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAG
GGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGACCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGACCTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTCCAG
CACAACTCCAGTGGGGCCGGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCAG
CAATTCTGAGTCCAGCACACCTCCAGTGGGGCCACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGG
GGCCAACACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAG
CACAACTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCACACAGCCACCAACTCTGGGTCCAG
TGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGAATGCACACAACTTCCCATAGTGATCTACTGCAGT
GAGTGAGGCAAAGCCTGGTGGGTCCCTGGTGCCGTGGGAAATCTTCCTCATCACCTGGTCTCGGTTGTGGCGGC
CGTGGGGCTCTTTGCTGGGCTCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGT
CTACCACCTCATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGAACAGCGGGCC
CTGAGCAGCCCCGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCACCCAAGACCTGGTTTCCTTT
CATTTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAAAATCTTGAAGAAGGTATTCTTCACCTTTCT
TGCCTTTACCAGACACTGGAAAGAGAATACTATATTGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACAC
ACACGACAAAGAGAAGCTGTGCTTGCCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAAC
CTCCATGCTGGACTCCATCTGGCATTCAAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

2025-09-23 10:55:00

312/615

FIGURE 310

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATIS
GSSVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESS
TPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTS
SGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGA
GTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTA
TNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNS
ESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESS
TVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGS
LVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLNTFNTAVYHPHGLNHGLGPGPGGNHGAP
HRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

20570 " 2355001

FIGURE 311A

CTAAGCCGGAGGATGTGCAGCTGCGGCGGCGCGCGCGCTACGAAGAGGACGGGGACAGGCGCCGTGCGAACCGA
 GCCCAGCCAGCCGGAGGACGCGGGCAGGGCAGGGACGGGAGCCCCGACTCGTCTGCCGCGCCCGTCTGCGCCGTCG
 TGCCGGCCCCCGCTCCCCGCGCGGAGCGGGAGGAGCCGCGCCACCTCGCGCCCCGAGCCGCGCTAGCGCGCGC
 CGGGCATGGTCCCCTCTTAAAGGCGCAGGCCGCGGCGGGGGCGGGTGTGCGGAACAAAGCGCCGGCGCGGGG
 CCTGCGGGCGGCTCGGGGGCCGCGATGGGCGCGGCGGGCCCCGCGGCGGCGGCGGCGCTGCCCGGGCGGGGCTCG
 CGGCGCTAGGGCGGGCTGGCCTCCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGGCGGGC
 GGCGGCGGGCGGGCGGGCGGGCGGGCGGAGCGGCGCGGGC**ATG**GCGCGCGCGGGCGGGCGCGCCTGGCTCAGCGTGC
 TGCTCGGGCTCGTCTGGGCTTCTGTGCTGGCCTCGCGGCTCGTCTGCCCGGGCTTCCGAGCTGAAGCGAGCGG
 GCCCAGGCGCGCGGCCAGCCCCGAGGGCTGCCGCTCCGGGCGAGGCGGCGGCTTCCAGGCGGGCGGGGCGCGCG
 GCGATGCGCGCGGGGCGCAGCTCTGGCCGCCCGGCTCGGACCCAGATGGCGGCCCCGCGGACAGGAACCTTCTCT
 TCGTGGGAGTCATGACCGCCAGAAATACCTGCAGACTCGGGCGCTGGCCGCTACAGAACATGGTCCAAGACAA
 TTCTGGGAAAGTTTCACTTCTTCTCAAGTGAGGGTTCTGACACATCTGTACCAATTCCAGTAGTGCCACTACGGG
 GTGTGGACGACTCCTACCCGCCCCAGAAGAAGTCTTTCATGATGCTCAAGTACATGCACGACCACTACTTGGACA
 AGTATGAATGGTTTATGAGAGCAGATGATGACGTGTACATCAAAGGAGACCGTCTGGAGAACTTCTGAGGAGTT
 TGAACAGCAGCGAGCCCCCTCTTCTTGGGCGAGCGGCTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
 AGCCTGGTGAGAACTTCTGCATGGGGGGGCTGGCGTGATCATGAGCCGGGAGGTGCTTTCGGAGAATGGTGCCGC
 ACATTGGCAAGTGTCTCCGGGAGATGTACACCACCCATGAGGACGTGGAGGTGGGAAGGTGTGTCCGGAGGTTTG
 CAGGGGTGCAGTGTGTCTGGTCTTATGAGATGCGGCGAGCTTTTTTATGAGAATTACGAGCAGAAACAAAAGGGGT
 ACATTAGAGATCTCCATAACAGTAAATTCACCAAGCTATCACATTACACCCCAACAAAAACCCACCTACAGT
 ACAGGCTCCACAGCTACATGCTGAGCCGCAAGATATCCGAGCTCCGCCATCGCACAAATACAGCTGCACCCGCAAA
 TTGTCTCTGATGAGCAAAATACAGCAACACAGAAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCTTCA
 TGAGGTTTCAGCCCCGCCAGCGAGAGGAGATTCTGGAATGGGAGTTTCTGACTGGAAAATACTTGTATTTCGGCAG
 TTGACGGCCAGCCCCCTCGAAGAGGAATGGACTCCGCCAGAGGGAAGCCTTGGACGACATTTGCATGCAGGTCA
 TGGAGATGATCAATGCCAACGCCAAGACCAGAGGGCGCATCATTGACTTCAAAGAGATCCAGTACGGCTACCGCC
 GGGTGAACCCCATGTATGGGGCTGAGTACATCCTGGACCTGCTGCTTCTGTACAAAAGCACAAAGGGAAGAAAA
 TGACGGTCCCTGTGAGGAGGACGCGTATTTACAGCAGACTTTTCAGCAAAATCCAGTTTGTGGAGCATGAGGAGC
 TGGATGCAAGAGTTGGCCAAAGAGAATCAATCAGGAATCTGGATCCTTGTCTTTCTCTCAAACTCCCTGAAGA
 AGCTCGTCCCCCTTTCAGCTCCCTGGGTGCAAGAGTGAGCACAAAGAACCCAAAGATAAAAAGATAAACATACTGA
 TTCTTTTGTCTGGGCGTTTCGACATGTTTGTGAGATTATGGGAAACTTTGAGAAGACGTGTCTTATCCCCAATC
 AGAACGTCAAGCTCGTGGTCTTGCTTTTCAATTCTGACTCCAACCCGTGACAAGGCCAAACAAGTTGAAGTATGA
 GAGATTACCGCATTAAGTACCCATAAGCCGACATGCAGATTTTGCCTGTGTCTGGAGAGTTTTCAGAGCCCTGG
 CCCTGGAAGTAGGATCCTCCAGTTTAAACAATGAATCTTGTCTCTTCTGCGACGTGCGACTCGTGTTTACTA
 CAGAATTCCTTCAGCGATGTGAGCAAAATACAGTTCTGGGCCAACAAATATATTTTCCAATCATCTTCAGCCAGT
 ATGACCCAAAGATTGTTTATAGTGGGAAAGTTCCAGTGACAACCATTTTGCCTTTACTCAGAAAAGTGGCTTCT
 GGAGAACTATGGGTTTGGCATCACGTGTATTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA
 TCCAAGGCTGGGGGCTGGAGGATGTGGACCTTTTCAACAAGGTTGTCCAGGCAGGTTTGAAGACGTTTAGGAGCC
 AGGAAGTAGGAGTAGTCCACGTCCACCATCTGTCTTTTGTGATCCCAATCTTGACCCCAACAGTACAAAATGT
 GCTTGGGGTCCAAGCATCGACCTATGGGTCCACCAGCAGCTGGCTGAGATGTGGCTGGAAAAAATGATCCAA
 GTTACAGTAAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCC**TAA**TGTCCAGCTTTGCTGGAAAAGACGTTTT
 TAATTATCTAATTTATTTTCAAAAATTTTTTGTATGATCAGTTTTTGAAGTCCGTATACAAGGATATATTTTAC
 AAGTGGTTTTCTTACATAGGACTCCTTTAAGATTGAGCTTTCTGAACAAGAAGGTGATCAGTGTTTGCCCTTTGAA
 CACATCTTCTTGCTGAACATTATGTAGCAGACCTGCTTAACCTTGAAGTTGAAATGTACCTGATGAACAAAACTTT
 TTTAAAAAATGTTTTCTTTTGGAGACCCTTTGCTCCAGTCTTATGGCAGAAAACGTGAACATCTCTGCAAGTAT
 TATTGTAACAAAACACTGTAACCTCTGGTAAATGTTTGTGTGATTGTTAACATTCCACAGATTCTACCTTTTGT
 GTTTTGTTTTTTTTTTTTTTACAAATGTTTTTAAAGCCATTTTCATGTTCCAGTTGTAAGATAAGGAAATGTGATAATA
 GCTGTTTCATCATTTGCTTTCAGGAGAGCTTTCCAGAGTTGATCATTTCTCTCATGGTACTCTGCTCAGCATGGC
 CACGTAGGTTTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG
 CAGTGGCGCAATCTTGGCTCACTTTAACCTCCACTTCCCTGGTTCAAGCAATCCCCTGCCCTTGGCTCCCCGAGT
 AGCTGGGATTACAGGCACACACCACCACGCCAGNATAGTTTTTTTGTATTTTTTAGTAGAGACGGGGTTTACCAT
 GCAAGCCCAGCTGGCCACGTAGGTTTTTAAAGCAAGGGCGGTGGAAGGACAGTGAAGGATGTGGCTGTTTCTCG
 TGGTAGTTTCACTCGGCCATAATAGACCTGGCATTAAATTTTCAAGAAGGATTTGGCATTTTCTCTTCTTGACCCCTT
 CTCTTTAAAGGGTAAAATATTAATGTTTAGAATGACAAAGATGAATTATTACAATAAATCTGATGTACACAGACT
 GAAACATACACACATACACCCTAATCAAAACGTTGGGGAAAAATGTATTTGGTTTTGTTTCTTTCATCCTGTCTG
 TGTATGTGGGTGGAGATGGTTTTTCATTCTTTCATTACTGTTTTGTTTTATCCTTTGTATCTGAAATACCTTTAA

FIGURE 311B

[illegible]

315/615

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

><subunit 1 of 1, 802 aa, 1 stop

><MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRAWLSVLLGLVLGFVLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGARG
DARGAQLWPPGSDPDGGPRDRNFLVGVMTAQKYLQTRAVAAAYRTWSKTIPGKVQFFSSEGS
TSVPIPVVPLRGVDDSYPPQKKSFMMLKYMHDHYLDKYEFMRADDDVYIKGDRLENFLRSLN
SSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGHKCLREMYTT
VEVGRCVRRFAGVQCVWSYEMRQLFYENYEQNKKGYYIRDLHNSKIHQAITLHPNKNPPYQYRL
HSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFPQPRQREEILEW
LTGKYLVSVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQYGYRRVNP
MYGAEYILDLLLLYKKHKGKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKRINQESGSL
SFLSNSLKKLVPFQLPGSKSEHKPKDKKINILIPLSGRFDMFVRFMGNFEKTCLIPNQNVKL
VVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQFNNESSLFFC
DVDLVFTTEFLQRCRANTVLGQQIYFPPIIFSQYDPKIVYSGKVPSDNHFQKTGFWRNYGF
GITCIYKGDIVRVGGFDVSIQGWGLEDDVLFNKVVQAGLKTFRSQEVGVVHVHHPVFCDPNLD
PKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

Signal peptide:

amino acids 1-23

Downloaded from www.nrj.com

316/615

FIGURE 313

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCC
TACGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACA
GGTGGGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCA
ACAAATGGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCAT
GGTCAATTAACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT
TTGGAATCATGGTGTGATGGAAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTT
TTGGAAGCATTTTTCATGCTGAGTCCCTTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATC
GCTGGATCAACAACCGCCTTGTGGCAACATGGCTCACCCCTACCTGTGGCATTATTGGAGACCA
TGTTTGGTGTAAGAGTGATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCAATTA
TCATGAACCATCGGACAAGAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCT
ACCTCAGATTGGAGAAAATTTGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGG
CCATGCAGGCTGCTGCCTATATCTTCATTCAATAGGAAATGGAAGGATGACAAGAGCCATTTTCG
AAGACATGATTGATTACTTTTGTGATATTCACGAACCACTTCAACTCCTCATATTCCCAGAAG
GGACTGATCTCACAGAAAACAGCAAGTCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTC
AGAAATATGAATATGTTTTACATCCAAGAACTACAGGCTTTACTTTTGTGGTAGACCGTCTAA
GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACACATTCCTC
AATCAGAGAAGCACCTCCTCCAAGGAGACTTTCCCAGGGAAATCCACTTTCACGTCCACCGGT
ATCCAATAGACACCTCCTCCACATCCAAGGAGGACCTTCAACTCTGGTGCCACAAACGGTGGG
AAGAGAAAGAAGAGAGGCTGCGTTCTTCTATCAAGGGGAGAAGAATTTTTATTTTACCGGAC
AGAGTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTCCTTGTGGTCAAATTGCTCTCTATAC
TGTATTGGACCCTGTTTACGCCCTGCAATGTGCCTACTCATATATTTGTACAGTCTTGTAAAGT
GGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGAGAATATTTGGTGGACTGGAGA
TCATAGAACTTGATGTTACCGACTTTTACACAAACAGCCACATTTAAATTCAAAGAAAAATG
AGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTGGAATGTTCTAAACCTT
TCTAAGCTCAGATGCATTTTGTGATGACTATGTGCAATATTTCTTACTGCCATCATTATTTGT
TAAAGATATTTTGCACCTTAATTTTGTGGGAAAAATATTGCTACAATTTTTTTTAAATCTCTGAA
TGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTTGGGCCAGAATA
TTATTAAACAATCATCAGGCTTTTAAA

317/615

FIGURE 314

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIFML
SPFLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTR
MDWMFLWNCLMRYSYLRLLEKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYF
CDIHEPLQLLIFPEGTDLTENSKSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLD
AVHDITVAYPHNIPQSEKHLLOGDFPREIHFHVHRYPIDTLPTSKEQLQWCHKRWEEKEERL
RSFYQGEKNFYFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITI
VIFVLQERIFGGLEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 315

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGA
 CCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGG
 AATATCCATGGCTTTTTGTGCTCATTTTTGGTTCTCAGTTTCTACGAGCTGGTGTGACGACAGTG
 GCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTG
 CTCCCTCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCA
 TGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCG
 AGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAAA
 CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTTCCAGATTTTACGATGAGGA
 GGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATA
 TGTTGACGGAGGTATCCAGTTACTCTGCCTGTCCTCAGGCTGGTTCCCCCAGCCACAGCCAA
 GTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAG
 CCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCAT
 CCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCA
 GCCCTCACCTTGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGT
 TGTGATGGGGATGATAATTGTTTTCTTCAAATCCAAGGGAAAATCCAGGCGGAACTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGAAACACGCAGTGGAGGTGACTCT
 GGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAA
 AGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCA
 GGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGT
 GGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACAATGTGACTTTGTCTCCCAACAATGG
 GTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTTACATTTCAATCCCCATTTTATCAG
 CCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTGGACTATGAGGGTGGGACCAT
 CTCCTTCTTCAATACAAATGACCAGTCCCTTATTTTATACCCTGCTGACATGTCAGTTTGAAGG
 CTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTCAT
 ATGTCCAGTGTCTTGGGGATTGAGACAGAGAAGACCCTGCTTAAAGGGCCCCACACCACAGACC
 CAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCTCCGGAGCCTGCGC
 ACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCCTGAGCCCTGCA
 GCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAGTCAGAAGCC
 ATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAACCTCCATC
 CAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGGACAGT
 GATTCTTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTTTGA
 GGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTTT
 GCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACAC
 CTGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTTCACTGATTCTATAAGCCCAG
 CATTACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCA
 TATCCCTCATTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAATTAACTAAACAAT
 ATATTTAAAGATGATATATAACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAA
 TATTTAAATATCAACCAGTGTAATTGAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAA
 AAAAAAAA

319/615

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866
><subunit 1 of 1, 466 aa, 1 stop
><MW: 52279, pI: 6.16, NX(S/T): 2
MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAV
VHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEAT
WELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLY
DVEISIIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVM
GMIIVFFKSKGKIQAELDWRKKGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDVRGKNNVTLSPNNGYW
VLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLL
RPYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

320/615

FIGURE 317

GCACCTGCGACCACCGTGAGCAGTCATGGCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTTC
TGGGCTTGTCTCTGGCTCTGTCGCTGCTGCTGCCCAAGGCCTTCCTGTCCCGCGGGAAGCGGCA
GGAGCCGCCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCACCA
GGCACCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGGCATT
TGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTGATGGG
GCAGATTATTCCAATCTACGGTTTTTGGGATTTTTTTTATATATACTGTACATTCTATTTAAGGT
AAGTAGAATCATCCTAATCATATTACATCAATTGAAAATCTAATATGGCGATAAAAATCATTGT
CTACATTAAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTAAATCCTGC
CTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCACTTTCACATAAGAATGTTTACTCAA
TGTTTAAGTGTTTTGCCCCAAAATTCACTAACAAGGCAGAACTAGGACTTGAACATGGAT
CTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

320/615

321/615

FIGURE 318

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPTPEGKLGRFPPMMHHHQAPSDGQTP
GARFQRSHLAEAFKAKGSGGGAGGGSGRGLMGQIIPIYGFGLFLYILYILFKVSRILIILHQ

FIGURE 319

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAG
 GAAAAGAGTTTGTGGGAACCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCT
 GGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTA
 CTATAGCACATTGTCATTTACAACACTGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAA
 CAATTTTACAGAAATGAGCCAGAGACTGAATCAATGGTGAAAAATGCATTTTATAAATCTCC
 ATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTT
 GGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAGATAAAAT
 TGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAGATCCTCA
 CTCAGTTAAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGG
 AACACGAAGAAGTAAACTCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGA
 GGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTT
 AATTAATGCCACATGGCTTGTGAGTGTGCTCACTGTTTACAACATATAAGAACCCTGCCAG
 ATGGACTGCTTCCTTTGGAGTAACAATAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAAT
 AATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCTCTTGACAGAGCTTTC
 TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGATGCATCCTATGAGTT
 TCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAATGATGGTTACAGTCA
 AAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACCTCAAGC
 TTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGC
 ATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTGAGATGCTAGAGATATCTGGTACCTTGC
 TGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTTTATACTAGAGT
 TACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAA
 CAGATAACATTTTTTTTTTGGTGTGGAGGCCATTTTTTAGAGATACAGAATTGGAGA
 AGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTTTGCTTGATGCATGTATTTT
 CTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATCAACTCTGTCATCTGTGA
 GCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATAACAATATTACATTACAGCCTGTA
 TTCATTTGTTCTCTAGAAGTTTTGTGACAGATTTTGAATTTGACTTGTGACATAAATTTGTAATGCATA
 TATACAATTTGAAGCACTCCTTTTCTTCAGTTCTCAGCTCCTCTCATTTTCAAGCAATATCCA
 TTTTCAAGGTGCAGAACAAAGGAGTGAAAGAAAAATATAAGAAGAAAAAATCCCCTACATTTTA
 TTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCATTAT
 GAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACATA
 GTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCAATTTCAAACAACACTACTATG
 ATAAATGTGAAGAAGATTCTGTTTTTTGTGACCTATAATAATTATACAACTTCATGCAATG
 TACTTGTTCTAAGCAAATTAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATA
 TAACAATAAAATATAAATCACCCA

323/615

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871

><subunit 1 of 1, 423 aa, 1 stop

><MW: 47696, pI: 8.96, NX(S/T): 3

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDK
LYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFH
STEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQS
LRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIK
PSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVYPTNAVHRVCLPDASYEFQPGDVMFVTG
FGALKNDGYSQNHRLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLV
SSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

323/615

FIGURE 321

CCGGGCTCCTCGGGTGAGGCCGGCAAGTTTGGAGAGCGTGGTCAGACAATAGGGGCGTGGCTACGG
CTCGCGGAGCGCAACCAACGCTCTAGACCAGACCTGGGCTCGAGACCATAACTGTTTGGCTTT
AACAGTACGTGGGCGGCCGGAATCCGGGAGTCCGGTGACCCGGGCTGTGGTCTAGCATAAAGG
CGGAGCCCAGAAGAAGGGGCGGGGTATGGGGAGAAGCCTCCCCACCTGCCCCCGCAAGGCGGCA
TCTGCTGGTCCTGCTGCTGCTCCTCTCTACCCTGGTGATCCCCTCCGCTGCAGCTCCTATCCA
TGATGCTGACGCCCAAGAGAGCTCCTTGGGTCTCACAGGCCTCCAGAGCCTACTCCAAGGCTT
CAGCCGACTTTTTCCTGAAAGGTAACCTGCTTCGGGGCATAGACAGCTTATTCTCTGCCCCCAT
GGACTTCCGGGGCCTCCCTGGGAACTACCACAAAGAGGAGAACCAGGAGCACCAGCTGGGGAA
CAACACCCTCTCCAGCCACCTCCAGATCGACAAGATGACCGACAACAAGACAGGAGAGGTGCT
GATCTCCGAGAATGTGGTGGCATCCATTCAACCAGCGGAGGGGAGCTTCGAGGGTGATTTGAA
GGTACCCAGGATGGAGGAGAAGGAGGCCCTGGTACCCATCCAGAAGGCCACGGACAGCTTCCAC
ACAGAACTCCATCCCCGGGTGGCCTTCTGGATCATTAAAGCTGCCACGGCGGAGGTCCCACCAG
GATGCCCTGGAGGGCGGCCACTGGCTCAGCGAGAAGCGACACCGCCTGCAGGCCATCCGGGAT
GGACTCCGCAAGGGGACCCACAAGGACGTCTAGAAAGAGGGGACCGAGAGCTCCTCCCACTCC
AGGCTGTCCCCCGAAAGACCCACTTACTGTACATCCTCAGGCCCTCTCGGCAGCTGTAGGGG
TGGGGACCGGGGAGCACCTGCCTGTAGCCCCCATCAGACCCTGCCCCAAGCACCATATGGAAA
TAAAGTTCTTTCTTACATCTAAAAA

325/615

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68879

><subunit 1 of 1, 242 aa, 1 stop

><MW: 27007, pI: 8.68, NX(S/T): 2

MGEASPPAPARRHLLVLLLLLSTLVIPSAAAPIHDADAQESSLGTLGLQSLQGFSLFL
KGNLLRGIDSLFSAPMDFRGLPGNYHKEENQEHQLGNNTLSSHLQIDKMTDNKTGEVLIS
ENVVASIQPAEGSFEGDLKVPRMEEKEALVPIQKATDSFHTELHPRVAFWIIKLPRRRSH
QDALEGGHWLSEKRHRLQAIRDGLRKGTHKDVLEEGTESSSHSRLSPRKTHLLYILRPSR
QL

Important features of the protein:

Signal peptide:

Amino acids 1-30

N-glycosylation sites:

Amino acids 97-101;112-116

N-myristoylation sites:

Amino acids 80-86;132-138;203-209;216-222

FIGURE 323

[illegible]

327/615

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

><subunit 1 of 1, 305 aa, 1 stop

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVV
HPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWS
DMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGC
SKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDR
REPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

20570" 6888001

FIGURE 325

AGCAGTGCATTGCTGGAGCGAGGAGAAGCTCACGAATCAGCTGCAGGTCTCTGTTTTGAAAAA
 GCAGAGATACAGAGGCAGAGGAAAAGGGTGGACTCCTATGTGACCTGTTCTTAGAGCAAGACA
 ATCACCATCTGAATTCCAGAAGCCCTGTTTCATGGTTGGGGATATTTTCTCGACTGCATGGAAAT
 CAGAAAAGAAGCAAAAGGATGGGAAATGCCTGCATTCCCCTGAAAAGAATTGCTTATTTCTCTAT
 GTCTCTTATCTGCGCTTTTTGCTGACTGAGGGGAAGAAACCAGCGAAGCCAAAATGCCCTGCCG
 TGTGTACTTGTACCAAAGATAATGCTTTATGTGAGAATGCCAGATCCATTCCACGCACCGTTC
 CTCCTGATGTTATCTCATTATCCTTTGTGAGATCTGGTTTTACTGAAATCTCAGAAGGGAGTT
 TTTTATTCACGCCATCGCTGCAGCTCTTGTTATTCACATCGAACTCCTTTGATGTGATCAGTG
 ATGATGCTTTTATTGGTCTTCCACATCTAGAGTATTTATTTCATAGAAAACAACAACATCAAGT
 CAATTTCAAGACATACTTTCCGGGGACTAAAGTCATTAATTCACCTTGAGCCTTGCAAACAACA
 ATCTCCAGACACTCCCAAAAGATATTTTCAAAGGCCTGGATTCTTTAACAAATGTGGACCTGA
 GGGGTAAATTCATTTAATTGTGACTGTAAACTGAAATGGCTAGTGGAATGGCTTGCCACACCA
 ATGCAACTGTTGAAGACATCTACTGCGAAGGCCCCCCAGAATACAAGAAGCGCAAAATCAATA
 GTCTCTCCTCGAAGGATTTTCGATTGCATCATTACAGAATTTGCAAAGTCTCAAGACCTGCCTT
 ATCAATCATTGTCCATAGACACTTTTTCTTATTTGAATGATGAGTATGTAGTCATCGCTCAGC
 CTTTTACTGGAAAATGCATTTTCCTTGAATGGGACCATGTGGAAAAGACCTTCCGGAATTATG
 ACAACATTACAGGCACATCCACTGTAGTATGCAAGCCTATAGTCATTGAAACTCAGCTCTATG
 TTATTGTGGCCAGCTGTTTGGTGGCTCTCACATCTATAAGCGAGACAGTTTTTGCAAATAAAT
 TCATAAAAATCCAGGATATTGAAATTCTCAAAATCCGAAAACCCAATGACATTGAAACATTCA
 AGATTGAAAACAACCTGGTACTTTGTTGTTGCTGACAGTTCAAAGCTGGTTTTACTACCATTTAC
 AAATGGAACGGAAACGGATTCTACTCCCATCAATCCTTACACGCGTGGTACAGGGACACTGAT
 GTGGAATATCTAGAAATAGTCAGAACACCTCAGACACTCAGAACGCCTCATTTAATTCTGTCT
 AGTAGTTCCCAGCGTCCTGTAAATTTATCAGTGGAACAAAGCAACACAATTATTCTACTAACCAA
 ACTGACATTCCTAACATGGAGGATGTGTACGCAAGTGAAGCACTTCTCAGTGAAAGGGGACGTG
 TACATTTGCTTGACAAGATTCATTGGTGATTCCAAAGTCATGAAATGGGGAGGCTCCTCGTTC
 CAGGATATTCAGAGGATGCCATCGCGAGGATCCATGGTGTTCAGCCTCTTCAAATAAATAAT
 TACCAATATGCAATTCTTGGAAGTGATTACTCCTTTACTCAAGTGTATAACTGGGATGCAGAG
 AAAGCCAAATTTGTGAAATTTTCAGGAATTAATGTTCAGGCACCAAGATCATTACACATGTG
 TCCATTAATAAGCGTAATTTTCTTTTTGCTTCCAGTTTTAAGGGAAATACACAGATTTACAAA
 CATGTCATAGTTGACTTAAGCGCATGAGACACCAAATTCTGTGGCTGCCATCAGAAATTTTCT
 ACAGTACATGACCCGGATGAACTCAATGCATGATGACTCTTCTTATCACACTTGCAAATGAAT
 GCCTTTCAAACATTGAGACTGCTAGAACCAAGCACTACCAGTATCTCCATCCTTAACTGTCCA
 GTCCAGTGATGTGGGAAGTTACCTTTATAAGACAAAATTTAATTGTGTAAGTGTCTTTGCA
 GTGAAGATGTGTAAATAAGCGTTTAATGGTATCTGTTACTCCAAAAAGAAATATTAATATGTA
 CTTTTCCATTTATTTATTTCATGTGTACAGAAACAACCTGCCAAATAAAATGTTTACATTTTCTT
 TCATA

329/615

FIGURE 326

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68882
><subunit 1 of 1, 557 aa, 1 stop
><MW: 63818, pI: 8.61, NX(S/T): 3
MESERSKRMGNACIPLKRIAYFLCLLSALLLLEGGKKPAKPKCPAVCTCTKDNALCENARS
IPRTVPPDVISLSFVRSGFTEISEGSFLFTPSLQLLLFTSNSFDVISDDAFIGLPHLEYL
FIENNNIKSISRHTFRGLKSLIHLNANNLQTLPKDIFKGLDSLNTNVDLRGNSFNCDCK
LKWLVEWLGHNTATVEDIYCEGPPEYKKRKINSLSKDFDCIITEFAKSQDLQPYQSLSID
TFSYLNDEYVIVIAQPFTGKCIFLEWDHVEKTFRNYDNITGTSTVVCKPIVIETQLYVIVA
QLFGGSHIYKRDSFANKFIKIQDIEILKIRKPNDIETFKIENNWFVADSSKAGFTTIY
KWNGNGFYSHQSLHAWYRDTDVEYLEIVRTPQTLRTPHLILSSSSQRPVIYQWNKATQLF
TNQTDIPNMEDVYAVKHFSVKGDVYICLTRFIGDSKVMKWGGSSFQDIQRMPSRGSVMVFQ
PLQINNYQYAILGSDYSFTQVYNWDAEKAKFVKFQELNVQAPRSFTHVSINKRNFLFASS
FKGNTQIYKHVIVDLA

Important features of the protein:

Signal peptide:

Amino acids 1-34

Transmembrane domain:

Amino acids 281-306

N-glycosylation sites:

Amino acids 192-196;277-281;422-426

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 310-314

Tyrosine kinase phosphorylation sites:

Amino acids 228-235;378-385

N-myristoylation sites:

Amino acids 172-178;493-499

Amidation site:

Amino acids 33-37

329/615

FIGURE 327

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCC
 CTACCAAGTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGA
 AAGAGAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTC
 CAGAAGGATGCCTCCATTTCCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACC
 CGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA
 CCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCA
 CTTACGCGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAAC
 CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCA
 GGCTTGTGCCAGCTTCAATGGGAACTGCTGTCTCTGGAACACCACGGTGGAAAGTCAAGGCTTG
 CCCTGGAGGCTACTATGTGTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGG
 TCATTTTTTATGACATCTGCGACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCGAGTGCAC
 ATGCGCTCCAGGAACTGTGCTAGGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGA
 GCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCCTACCGCTGTGAGTG
 TGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAA
 TAACAATGGTGGCTGCAGCCACTCTTGCCTTGATCTGAGAAAGGCTACCAGTGTGAATGTCC
 CCGGGGCCTGGTGTCTGTGAGGATAACCACACTTGGCCAAGTCCCTGTGTTGTGCAATCAAA
 TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTTGGTGGCCTGGAGCTCTTCCTGACCAACAC
 CTCCTGCCGAGGAGTGTCCAACGGCACCACATGTCAACATCCTCTTCTCTCTCAAGACATGTGG
 TACAGTGGTTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGACAGGTCTACCCAA
 GCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTGAC
 CTGCGAGTTTCCACGCCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCC
 ACTGGAATCATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAA
 TGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCACCCTCAAGCTTCGTGACTCCCTCTACTT
 TGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGGAAGCTTGGTGGAGAGCTGCTTTGCCAC
 CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGATGGCTGTGTTTCAGA
 TGAATCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGGTCCCTGTCTT
 CAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGT
 GTTGGACGAGCGTTCCCGCTGTGCCCAGGGTTGCCACCGGCGAATGCGTCGTGGGGCAGGAGG
 AGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGA
 GGACTTAGTTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTC
 CCCCCACCGCCCTCTAAGAACATCTGCCAACAGCTGGGTTTCAGACTTCACACTGTGAGTTCAG
 ACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCAAGTGGGCACAGGTACAGCACTGCTG
 AACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAACTGTCCACCCAGAA
 AGACACTCACCCCATTTCCCTCATTTCTTTCTTACACTTAAATACCTCGTGTATGGTGCAATC
 AGACCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAG
 TTAATGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTT
 GAAATTTCAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGTATCAATAAAACAGTAT
 ATAATTTT

331/615

FIGURE 328

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLQDESQGPPLCDNHVNGEWYHFT
GMAGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPG
GYVYRLTKPSVCFHVYCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQN
NGGCSEICVNLKNSYRCECGVGRVLRSDGKTCEDEVEGCHNNNGGCSHSCLGSEKGYQCECPRG
LVLSEDNHTCQVPVLCKSNAIEVNIPRELVGGLFLTNNTSCRGVSNNGTHVNI LFSLKTCGTV
VDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLE
IMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLRLDSLYFGIEPVVHVSGLES LVESCFATPT
SKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLVCGVLD
ERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 329

GAGAGAGGCACGACGACTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCT
GCATCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACC
TGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACC
GGGAGGATCACAGAGCCAGCATGTTACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG
ATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCA
TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCTCATCAAGGTGATTG
TGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTG
ACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAG
GGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA
CAGGGAACTGGTTCTCTGCCTGTTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTA
GGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAA
TCACAGAAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCC
TGGTCTCCCTGCACTGTCTTGCCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGG
AGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAGTACGACAAACAGCACGTCT
GTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAACATA
CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAGCTTCCCATCCCTGG
CTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACATCGCCC
TCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCT
TTGATGAGGAGCTCACTCCAGCCACCCCCACTCTGGATCATTTGGATGGGGCTTTACGAAGCAGA
ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGT
GCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCCG
AAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCAGT
GGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGCCCGAGCACCCCAGGAGTAT
ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTGTAATGTCT
GCTGCCCCCTTTGCAGTGCTGGGAGCCGCTTTCCTTCTGCCCCTGCCACCTGGGGATCCCCCAA
AGTCAGACACAGAGCAAGAGTCCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCATTTCTT
GGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCCTCGCAGCCCAGAGGCGCCCAGAGGAAG
TCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACT
GAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCACACTACTGAATGGAAGC
AGGCTGTCTTGTAAGAGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCA
GCCCTGTCCGTCTTCACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT
GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTATTGTTATTACAGCTATGG
CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAAA

333/615

FIGURE 330

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47644, pI: 5.18, NX(S/T): 2
MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIVILDKYYFL
CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSA
CFDNFTEALAEACRQMGYSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCL
ACGKSLKTPRVVGGEESVDSWPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDVFNWK
VRAGSDKLGSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTP
ATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTTC
QGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEI
```

Transmembrane domain:
amino acids 32-53 (typeII)

005336 0150

FIGURE 331

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGGC
TCCAGGACTTTGGCCATCTATAAAGCTTGGCA**ATG**AGAAATAAGAAAATTCTCAAGGAGGACG
AGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTCGAAATCA
ATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATCTACC
TGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGCGCGGC
TCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGTCCTTCT
CCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTGCAAGTCC
TGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGTAGACAACCT
TCACTCAGAACCCAGGGATGTTTCAAGATCAAAGGTGAACAAGGCGCCCCAGGTCTTCAAGGTC
ACAAGGGGGCCATGGGCATGCCTGGTGGCCCTGGCCCGCCGGGACCACTGCTGAGAAGGGAG
CCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGACCCCAAGGCCACCGGGAG
TCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAGGAGCCACTGGCA
CCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATTTGGCCCCAAAAGGGG
AAACTGGAACCTAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAAAGGGGACAGGGGCA
TGAAAGGAGATGCAGGGGTCTATGGGGCCTCCTGGAGCCCAGGGGAGTAAAGGTGACTTCGGGA
GGCCAGGCCACCAGGTTTGGCTGGTTTTCTGGAGCTAAAGGAGATCAAGGACAACCTGGAC
TGCAGGGTGTTCCGGGCCCTCCTGGTGCAGTGGGACACCCAGGTGCCAAGGGTGAGCCTGGCA
GTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGAGTCCAGGAGCCACAGGCC
TGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGAAGAAAAGGAGAATCAGGAG
TTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGGGCTGGCAGGTCCCAAGGGAG
CCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAGGATCTTCTGGGGAGCAAGGAG
TAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACTCAGTGTCCGTCAGGATTGTCGGCAGTAGTA
ACCGAGGCCCGGGCTGAAGTTTACTACAGTGGTACCTGGGGGACAATTTGCGATGACGAGTGGC
AAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTACTCCAAAGGAAGGGCCCTGTACA
AAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAATGTTTCAAGTGTGCGGGCACGGAGAGTA
CCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCATGACTGCAGCCACGAGGAGGACGCAG
GCGTGGAGTGCAGCGTCT**TGA**CCCGGAAACCTTTCACTTCTCTGCTCCCGAGGTGTCTCTGGG
CTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGTTCCTGGGGACAACCTGAGCAGCCTCTGG
AGAGGGGCCATTAAATAAAGCTCAACATCATTGA

335/615

FIGURE 332

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886
><subunit 1 of 1, 520 aa, 1 stop
><MW: 52658, pI: 9.16, NX(S/T): 3
MRNKKILKEDELLSETQQAAPHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGLL
VVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVRVS
HEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRDGAT
GPSGPQGPVGKGEAGLQGPQGA PGKQGATGTPGPQGEKGSKG DGGLIGPKGETGTKGEKDDL
GLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFP GAKGDQGPGLQGVPGPPGAV
GHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGGQGRKGESGVPGPAGVKGEQ
GSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGRAEVYYSG
TWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLWSCTKNSWG
HHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 333

GGGCTGTTGATTTGTGGGGGATTTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCTG
 CCTCTGGCATATGCACACACTCACACATTCTGTACACCCGTCACACACACATACCATGTTCT
 CCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGCTACCCTGAAGCTCTGGC
 TGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCAACTG
 CTGTCACTGCATGCTCTGCCAAGGAGGAGGGAAGTGCAGTGACAGCAGGAGTAAGAGTGGGAG
 GCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTGAGCGAGCCTAGAGAGGGCAGACTA
 TCAGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGGCAGAAGACCGG
 GGCACCTTGTGGGTTGCAGAGCCCTCAGCC**ATG**TTGGGAGCCAAGCCACACTGGCTACCAGGT
 CCCCTACACAGTCCCGGGCTGCCCTTGTTCTGGTGCTTCTGGCCCTGGGGGCCGGGTGGGCC
 CAGGAGGGGTGAGAGCCCGTCTGTGAGGGGGAGTGCCCTGGTGGTCTGTGAGCCTGGCCGA
 GCTGCTGCAGGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCTGGGCGAGTGGCATT
 GCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGCACCAGTGGGGCC
 ATCTACTTCGACCAGGTCTGTGTAACGAGGGCGGTGGCTTTGACCGGCCTCTGGCTCCTTC
 GTAGCCCTGTCCGGGGTGTCTACAGCTTCCGGTTCCATGTGGTGAAGGTGTACAACCGCCAA
 ACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCTCAGCCTTTGCCAATGATCCT
 GACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGACCCCTGGGGACCGAGTG
 TCTCTGCGCCTGCGTCCGGGGAAATCTACTGGGTGGTTGGAAATACTCAAGTTTCTCTGGCTTC
 CTCATCTTCCCTCTCT**TGAGG**ACCCAAGTCTTTCAAGCACAGAATCCAGCCCTGACAACCTTT
 CTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAGACTCCCTCTGGCTCCTATC
 CCACCTCTTTGCATGGGACCCTGTGCCAAACACCCAAGTTTAAGAGAAGAGTAGAGCTGTGGC
 ATCTCCAGACCAGGCCTTTTCCACCCACCCACCCCAAGTTTACCCTCCAGCCACCTGCTGCATC
 TGTTCTCTGCCTGCAGCCCTAGGATCAGGGCAAGGTTTGGCAAGAAGGAAGATCTGCACTACTT
 TGCGGCCTCTGCTCCTCCGGTTCCCCCACCCAGCTTCTGCTCAATGCTGATCAGGGACAGG
 TGGCGCAGGTGAGCCTGACAGGCCCCACAGGAGCCAGATGGACAAGCCTCAGCGTACCCTG
 CAGGCTTCTTCTGTGAGGAAAGCCAGCATCACGGATCTCAGCCAGCACCGTCAGAAGCTGAG
 CCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCACAGGCAGAAGGGTGGGAAGGGCCTGGA
 GTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTGTCTAGACTGAACATGGTACACATTCTG
 CATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCTGGCTGTCCTTCTATGCTGGATCCCAGAT
 GGACTCTGGCCCTTACCTCCCCACCTGAGATTAGGGTGAGTGTGTTTGCTCTGGCTGAGAGCA
 GAGCTGAGAGCAGGTATACAGAGCTGGAAGTGGACCATGGAAAACATCGATAACCATGCATCC
 TCTTGCTTGGCCACCTCCTGAAACTGCTCCACCTTTGAAGTTTGAACCTTTAGTCCCTCCACAC
 TCTGACTGCTGCCTCCTTCTCCAGCTCTCTCACTGAGTTATCTTCACTGTACCTGTTCCAG
 CATATCCCCACTATCTCTCTTTCTCCTGATCTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCT
 ATTACCTGGGATTCCATGATTCAATCCTTCAGACCTCTCCTGCCAGTATGCTAAACCCTCCC
 TCTCTCTTTCTTATCCCGCTGTCCCATTTGGCCCAGCCTGGATGAATCTATCAATAAAACAACT
 AGAGAATGGTGGTCAGTGAGACACTATAGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGGA
 TCGGGTGTTACAGGTACAAGTAGGTATGTTGCAGAGGAAAATAAATATCAAACCTGTATACTAA
 AATTAAAAA

337/615

FIGURE 334

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180
><subunit 1 of 1, 205 aa, 1 stop
><MW: 21521, pI: 7.07, NX(S/T): 1
MLGAKPHWLPGPLHSPGLPLVLVLLALGAGWAQEGSEPVLLLEGECLVVCEPGRAAAGGPGGAA
LGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVYSF
RFHVVKVYNRQTVQVSLMLNTWVPVISAFANDPDVTREAATSSVLLPLDPGDRVSLRLRRGNLL
GGWKYSSFSGFLIFPL
```

Signal peptide:
amino acids 1-32

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 335

GCTGTTTCTCTCGCGCCACCACTGGCCGCCGGCCGCAGCTCCAGGTGTCCTAGCCGCCAGCC
TCGACGCCGTCCCGGGACCCCTGTGCTCTGCGCGAAGCCCTGGCCCCGGGGCCGGGGCATGG
GCCAGGGGCGCGGGGTGAAGCGGCTTCCCGCGGGGCCGTGACTGGGCGGGCTTCAGCCATGAA
GACCCTCATAGCCGCCTACTCCGGGGTCTGCGCGGCGAGCGTCAGGCCGAGGCTGACCGGAG
CCAGCGCTCTCACGGAGGACCTGCGCTGTGCGCGGAGGGTCTGGGAGATGGGGCACTGGATC
CAGCATCCTCTCCGCCCTCCAGGACCTCTTCTGTGTCACCTGGCTCAATAGGTCCAAGGTGGA
AAAGCAGCTACAGGTCACTCAGTGCTCCAGTGGGTCTGTCTTCTTGTACTGGGAGTGGC
CTGCAGTGCCATCCTCATGTACATATTCTGCACTGATTGCTGGCTCATCGCTGTGCTCTACTT
CACTTGGCTGGTGTGTTGACTGGAACACACCCCAAGAAAGGTGGCAGGAGGTACAGTGGGTCCG
AAACTGGGCTGTGTGGCGTACTTTTCGAGACTACTTTCCCATCCAGCTGGTGAAGACACACAA
CCTGCTGACCACCAGGAATAATCTTTGGATAACCACCCCATGGTATCATGGGCCTGGGTGC
CTTCTGCAACTTCAGCACAGAGGCCACAGAAGTGAGCAAGAAGTTCCAGGCATACGGCCTTA
CCTGGCTACACTGGCAGGCAACTTCCGAATGCCTGTGTTGAGGGAGTACCTGATGTCTGGAGG
TATCTGCCCTGTGAGCCGGGACACCATAGACTATTTGCTTTCAAAGAATGGGAGTGGCAATGC
TATCATCATCGTGGTTCGGGGGTGCGGCTGAGTCTCTGAGCTCCATGCCTGGCAAGAATGCAGT
CACCTGCGGAACCGCAAGGGCTTTGTGAACTGGCCCTGCGTCATGGAGCTGACCTGGTTCC
CATCTACTCCTTTGGAGAGAATGAAGTGTACAAGCAGGTGATCTTCGAGGAGGGCTCCTGGGG
CCGATGGGTCCAGAAGAAGTTCCAGAAATACATTGGTTTCGCCCCATGCATCTTCCATGGTTCG
AGGCCTCTTCTCCTCCGACACCTGGGGGTGGTGGCCCTACTCCAAGCCCATCACCACTGTTGT
GGGAGAGCCCATCACCATCCCCAAGCTGGAGCACCCAACCCAGCAAGACATCGACCTGTACCA
CACCATGTACATGGAGGCCCTGGTGAAGCTCTTCGACAAGCACAAGACCAAGTTCGGCCTCCC
GGAGACTGAGGTCTGGAGGTGAACTGAGCCAGCCTTCGGGGCCAATTCCTGGAGGAACCAG
CTGCAAATCACTTTTTTGCTCTGTAAATTTGGAAGTGTGTCATGGGTGTCTGTGGGTATTTAAA
AGAAATTATAACAATTTTGCTAAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

339/615

FIGURE 336

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184
><subunit 1 of 1, 388 aa, 1 stop
><MW: 43831, pI: 9.64, NX(S/T): 3
MKTLLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRSK
VEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLIADVLYFTWLVFDWNTPKKGRRSQW
VRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAFCNFSTEATEVSKKFPGIR
PYLATLAGNFRMPVLREYLMSSGICPVSRDTIDYLLSKNGSGNAIIIVVGAAESLSSMPGKN
AVTLRNRKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKKFQKYIGFAPCIFH
GRGLFSSDTWGLVPYSPKITT VVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDKHKTKFG
LPETEVLEVN

Important features of the protein:

Transmembrane domain:

amino acids 76-97

N-glycosylation sites.

amino acids 60-63, 173-176, 228-231

N-myristoylation sites.

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-245, 318-323, 378-383

FIGURE 337

GGGCGGCGGGATGGGGGCGGGGGCGGGCGGGCGCCGCACTCGCTGAGGCCCCGACGCAGGGCCGGGCGGGCCCCA
 GGGCCGAGGAGCGCGGCGGCCAGAGCGGGGCGCGGAGGCGACGCCGGGGACGCCCGCGCAGCAGGAGGTGGCG
 GCGGCTGCAGGCTTGTCCAGCCGGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGA
 CGGCTGTCTCAGCGAGGGGCGGTGCACCCGCTCCTGAGCAGCGCC**ATG**GGCCTGCTGGCCTTCTGAAGACCCA
 GTTCGTGCTGCACCTGCTGGTTCGGCTTTGTCTTCGTGGTGAGTGGTCTGGTCATCAACTTCGTCCAGCTGTGCAC
 GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCTCAACTGCCGCTCGCCTACTCACTCTGGAGCCA
 ACTGGTCATGCTGCTGGAGTGGTGGTCCTGCACGGAGTGACACTGTTACGGACCAGGCCACGGTAGAGCGCTT
 TGGGAAGGAGCACGCAGTCATCATCCTCAACCACAACCTTCGAGATCGACTTCTCTGTGGGTGGACCATGTGTGA
 GCGCTTCGGAGTGCTGGGGAGCTCCAAGGTCCTCGCTAAGAAGGAGCTGCTCTACGTGCCCCCTCATCGGCTGGAC
 GTGGTACTTTCTGGAGATTGTCTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTGAAGGGCTGAG
 GCGCTGTGCGACTACCCCGAGTACATGTGGTTCCTCTGTACTGCGAGGGGACGCGCTTCACGGAGACCAAGCA
 CCGCGTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCTGTCTCAAGTACCACCTGCTGCCGCGGACCAAGGG
 CTTACACCACCGCAGTCAAGTGCTCCGGGGGACAGTCGCAGCTGTCTATGATGTAACCTGAACCTCAGAGGAAA
 CAAGAACCCGTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGGCGGACATGTGCGTGAGGAGATTTCTCT
 GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA
 GGAGATATATAATCAGAAGGGCATGTTTCCAGGGGAGCAGTTTAAAGCCTGCCCGAGGCCGTGGACCCCTCTGAA
 CTTCTGTCTGGGCCACCATTTCTCTGTCTCCCTCTTCAGTTTTGTCTTGGGCGTCTTTGCCAGCGGATCACC
 TCTCCTGATCCTGACTTTCTTGGGGTTTGTGGGAGCAGCTTCTTTGGAGTTTCGCAGACTGATAGGAGAATCGCT
 TGAACCTGGGAGGTGGAGATTGCAGT**GA**GCTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT
 CAGTCTCAAAAAAAAAAAAAAAAAAACAACCCAGAAATTTCTGGAGTTGAACGTGTGTAGTTACTTGACATGAAAA
 ATTCACCTAGAGGCTGAACAGCAGATTTGAGCAGGCAGAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT
 TTTAGGCTAATGAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA
 CATATGCATGATGAGAGTCCAGAAAGGAGAGGAGAGAAAGGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA
 GTAACCTACCCACTCAGGAAGCTCAGTGAATCCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTTCAT
 AATCAAAGTGTCAAATGACAAAGAATCTTGAAGCAGCAAGAGATGAGCAACTTATCTTGTCAAAGGATCTTTG
 ATCAGATTAAACAGCTCATTTCTCCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA
 AACCTTCAACTGTAATTATTGGACTTTTGTAGTCTTAGATGGTCTGACCTCTTTGTCTTCAGGGACAGTTTTTCA
 ATTTAATCCCTAATAACAATTAGTCAAGCTTCTTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC
 TTACACCTGTAATCCAGCACTTTGGGAGGCCAGACGGGTGGATCATTTGGGGTCAGGCTGATCTCAAATCCT
 GAGTTACAGGTGATCTGCCCCGCTCAGCCTCCCAAAGTGTGTGATTGCAGGCGTGAGCCACTGCGCTGGCCGGA
 ATTTCTTTTTAAGGCTGAATGATGGGGGCCAGGCACGATGGCTCACGCCTGTGATCCCAAGTAGCTTTGGATTGTA
 AACATGCACCACCATGCGCTGGCTAATTTTTGTATTTTTTAGTAGAGACGTGTTAGCCAGGCTGGTCTCGATCTCCT
 GACCTCAAGTGACCACCTGCGCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA
 GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGGAAATGTCTGTTCAAGTCCTTTG
 CCTTTTTAAATTTTTATTATTTATTTATTTATTTTGGAGACAGGGTCTTGTCTGTGCCCCAGGCTGGAGTA
 CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCACCTCAGCCTCCCTTGT
 AGCTGTATTTTTTTGTATTTTGTATTTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGCATTTACCATGA
 TGCCAGGCTGGTCTTGAACCTCTGAGCTCAAGTGATCTGCCTGCTTCAGCCTCCCAAAGTGTGGGATTACAGA
 CATGAGCCACTGCACCTGGCAAACCTCCCAAATTAACACACACACAAAAAACCACTGATTCAAAATGGGCA
 GAGGGGCCGGGTGTGGCCCACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCGAGGCTG
 CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC

341/615

FIGURE 338

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213
><subunit 1 of 1, 368 aa, 1 stop
><MW: 42550, pI: 9.11, NX(S/T): 1
MGLLAFLKTQFVLHLLVGFVVFVSGLVINFVQLCTLALWPVSKQLYRRLNCRLAYSLWSQLVM
LLEWWSCTECTLFSTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKKEL
LYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHRVSM
EVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVYDVTNLNFRGNKNPSLLGILYGKKYEAD
MCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLLNFLSWA
TILLSPLFSFVLGVFASGSPLILTLFLGFVGAASFGVRRRLIGESLEPGRWRLQ
```

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 339

GATATTCTTTATTTTTAAGAATCTGAAGTACT**ATG**CATCACTCCCTCCAATGTCCTGGGGCAG
CCACCAGGCATATTCATCTTTGTGTGTGTTTTCTTTTGCTTTAGCACTGGGGCACTTCTTGC
TTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCGGCT
TACGCCTGATACGGCCCTGGGTTAGAAGGGAAGGGAAGATAAACTTTTATACAAATGGGGATA
GCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCCTGGGATCTGCCTATACCTTCTTTTCTC
TAACCTGGCATAACCCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAAAGTATT
TAGAGCTACAAGAGCCCTCATGGTCTGGCCCCCTGCCCCCCTGGCCAGCTTCATTGTACATGTG
GTGTTCTCTTGTCGTTCCCT**TAAT**GTGGTATGCCATGGGGTCTTTGCACAAGCCTTTCTCTT
TGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTCATCCTGCAG
ATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCATCTTGTTTAAT
GCTCTCATAAGACCACTTGTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTATCTTTATGTGC
GTTTGTGGTTGTATGGGTTGTGTCTGTTCCCCAGAATGCCAGCTCTGAGCTGCGTGAGGGTC
AAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCTCATGTTTTAGAGA
CTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCAGATGGTGTAGGGCC
CAGCATTGTAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGATGCAGGTCCTGATTCT
AGTAGGCCCAGGTTGGGCATCTCTAACAACTCCACGTGATGCTGATGCTGGTCCTATGAAC
TATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGCTCACACCTATGATCCCA
GCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATTTCAAGACTAGCCTGGCCA
ACATGGTGGAACCCCATCTGTACTAAAAATACACAAATTAGCTGGGCATGGTGGCACATGCCT
GTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTGAACCTGGGAGGCGGAGGTTG
CAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGACAGAGTGAGACTCTATGTCCAA
AAAAAAAAAAAA

343/615

FIGURE 340

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234
><subunit 1 of 1, 143 aa, 1 stop
><MW: 15624, pI: 9.58, NX(S/T): 0
MHHS LQCPGAATRHHLCVCF SFALALGHFLLISLVGKGLSLSCGVGGRQAGLRRLIRPWVRE
GKINFYTN GDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFS LFLGSKYLELQEPSWSGP
CPPGQLHCTCGVLLSFL
```

Important features of the protein:

Signal peptide:

amino acids 1-28

202110 030000T

FIGURE 341

CGCCATGGCCCGGGCTATCCCGCGGGTCCGCGCGCGCACTGCTCGCCGCCCTGCTGGCGTCGACG
CTGTTGGCGCTGCTCGTGTGCGCCCGCGCGGGGTGCGGGCGGCCGGGACCACGGGGACTGGGAC
GAGGCCTCCCGGCTGCCGCCGCTACCACCCCGCGAGGACGCGGGCGCGCGTGGCCCGCTTCGTG
ACGCACGTCTCCGACTGGGGCGCTCTGGCCACCATCTCCACGCTGGAGGCGGTGCGCGGCCGG
CCCTTCGCCGACGTCTCTCGCTCAGCGACGGGCCCCCGGGCGCGGGCAGCGGCGTGCCCTAT
TTCTACCTGAGCCCGCTGCAGCTCTCCGTGAGCAACCTGCAGGAGAATCCATATGCTACACTG
ACCATGACTTTGGCACAGACCAACTTCTGCAAGAAACATGGATTTGATCCACAAAGTCCCCTT
TGTGTTACATAATGCTGTCAGGAACTGTGACCAAGGTGAATGAAACAGAAATGGATATTGCA
AAGCATTCGTTATTTCATTCGACACCCTGAGATGAAAACCTGGCCTTCCAGCCATAATTGGTTC
TTTGCTAAGTTGAATATAACCAATATCTGGGTCCTGGACTACTTTGGTGGACCAAAAATCGTG
ACACCAGAAGAATATTATAATGTCACAGTTCAGTGAAGCAGACTGTGGTGAATTTAGCAACAC
TTATGAAGTTTCTTAAAGTGGCTCATAACACTTAAAGGCTTAATGTTTCTCTGGAAAGCGT
CCCAGAATATTAGCCAGTTTTCTGTC

345/615

FIGURE 342

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71269
><subunit 1 of 1, 220 aa, 1 stop
><MW: 24075, pI: 7.67, NX(S/T): 3
MAGLSRGSARALLAALLASTLLALLVSPARGRGGRDHGDWDEASRLPPLPPREDAARVAR
FVTHVSDWGALATISTLEAVRGRPFADVLSLSDGPPGAGSGVPYFYLSPLQLSVSNLQEN
PYATLTMTLAQTNFCKKHGFD PQSPLCVHIMLSGTVTKVNETEMDIAKHSLFIRHPEMKT
WPSSHNWFFAKLNITNIWVLDYFGGPKIVTPEEYYNVTVQ
```

Important features of the protein:

Transmembrane domain:

Amino acids 11-29

N-glycosylation sites:

Amino acids 160-164;193-197;216-220

N-myristoylation sites:

Amino acids 3-9;7-13;69-75;97-103

20250310 10:50:00

FIGURE 343

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGAT
TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTTCAGCAACTAAAAAAGCCACAG
GAGTTGAACTGCTAGGATTCTGACT**ATG**CTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT
AAAATCTGTTTTTTTGTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACT
GTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTTCTACCCACACCGT
CCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAAC
TGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGG
CTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGAT**TGA**CAAGCCCGAAGATTTTCATAGGCG
ATGGCTCCCCTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGG
GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTT
CATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCAAGGAACCCAGAGCT
TGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGA
CTTGCAATTCTCCTGGAACATGAGGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTG
TGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGC
ATTGAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACTCGGGACTCACCTCTG
GGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTGAAGGGCAACTGCAGGC
CGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAGCCAGGGGCAGCCGTCTG
GGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTCTGAGAGGCCCTCCTATGT
CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATGGCTCAGTGTTGGCCCAGGA
GGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCGAACACGGAAATGCCTCCAGT
AAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTTAAATCATGTTCTAGT
AATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTTCAAATGATCTCCAAGGGCCCT
TATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCAAACCTAAGAACCAGGTGCA
TTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTTTGGGAGGCCGAGGCG
GGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCTGTCTC
TACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATCCAGCTACTCGGG
AGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT
TCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAAGAATTATGGTTATTT
GTAA

347/615

FIGURE 344

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277
><subunit 1 of 1, 109 aa, 1 stop
><MW: 11822, pI: 8.63, NX(S/T): 0
MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTAS
PCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD
```

Signal peptide:
amino acids 1-15

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

349/615

FIGURE 346

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

><subunit 1 of 1, 671 aa, 1 stop

><MW: 74317, pI: 7.61, NX(S/T): 0

MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGPKDLFPYDK
CKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDERGV
AVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSEEN
SESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPVAMAR
SASSSSSSSSSSSDSDSVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEVDRISE
WKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGDELREDD
EPVKKRGRKGRGRGPPSSSDSEPEAELEAREAKSAKKPQSSSTEPARKPGQKEKVRPEEKQQ
AKPVKVERTRKRSEGF SMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDSPDVKRCLNALEELG
TLQVTSQILQKNTDVVATLKKIRRYKANKDVMKAAEVYTRLKSRVLGPKIEAVQKVNKAGME
KEKAEKLAGEEELAGEEAPQEKAEDKPSTDL SAPVNGEATSQKGESAEDKEHEEGRDSEEGPR
CGSSEDLHDSVREGPDLDRPGSDRQERERARGDSEALDEES

Signal peptide:

amino acids 1-13

TTTGGTTCTCCTGGATCTTTACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTTAAATTA
ACACCATTTGAAAGAGAACATTGTTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA
TTTTACCTTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTTCTTCCAGTCACTGGCACTTTGAAGCA
AAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCCTTTTTGGGTTTCATC
AGAAGGACTGGATTTTTCAAACCTCTTCTCTTAGATGAGGAAAGGAGGCAGGCTGCTCTGGGAGCCAAAGACCACAT
CTTTCTACTCAGTCTGGTTGACTTTAAACAAAATTTTAAAGAAATTTATTGGCCTGCTGCAAAGGAACGGGTGGA
ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTTTCATCAGAGTACTTCAGCCCTATAACAA
AACTCACATATATGTGTGTGGAACCTGGAGCATTTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA
GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTTCGATCCTCAGCAGCC
TTTTGCTTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTTCTTGGCAAAGATACTGCATT
CACTCGATCCCTTGGGCCCTACTCATGACCACACATACATCAGAACTGACATTTTCAGAGCTACTTGGCTCAATGG
AGCAAAATTTATTGGAACCTTTCTTCATACACGACACCTACAATCCAGATGATGATAAAATATATTTTCTTCTTTCG
TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAGAATGATGT
AGGAGGACAACGCAGCCTGATAAAACAAGTGGACGACTTTTTCTTAAAGCCAGACTGATTTGCTCAATTCTTGGGAAG
TGATGGGGCAGATACTTACTTTTGATGAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
AGTATATGGAGTCTTTACTACAACCAAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT
CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAGTGCAGACCATTCTTTGGGTGCAGTATGATGGGGAAT
TCCTTATCCACGGCCTGGTACATGTGCCAAGCAAAACCTATGACCACCTGATTAAGTCCACCCGAGATTTTTCCAGA
TGATGTCATCAGTTTCATAAAGCGGCACCTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCAACGTT
CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTCATTGCAGAAGATGGCCAGTACGA
TGTAATGTTTTCTTGGAACAGACATTGGAACCTGTCCTCAAAGTTGTTCAGCATTTCAAAGGAAAAGTGGAAATATGGA
AGAGGTAGTGTCTGGAGGAGTTTCAGACATATTCAGCACTCATCAATCATCTTGAACATGGAATGTCTCTGAAGCA
GCAACAATTGTACATTTGGTTCCGAGATGGATTAGTTTTCAGCTCTCTTGCACAGATGCGACACTTATGGGAAAGC
TGTCGCAGACTGTTGTCTTGGCAGAGACCCCTACTGTGCCTGGGATGGAAATGCATGCTCTCGATATGCTCCTAC
TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAATATGGCGACCCAATCACCCAGTGTCTGGGACATCGAAGACAG
CATTAGTCATGAACTGCTGATGAAAAGGTGATTTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC
TAAATCCCACAAAGCAACTATTAATGGTATATCCAGAGGTCAGGGGATGAGCATCGAGAGGAGTTGAAGCCCGA
TGAAAGAATCATCAAAACGGAATATGGCTACTGATTGAAAGTTGCAGAAGAAGGATTCTGGGATGTATTACTG
CAAAGCCAGGAGCACATTTTCATCCACACCATAGTGAAGCTGACTTTGAATGTCAATTGGAATGAACAGATGGA
AAATACCCAGAGGGCAGAGCATGAGGAGGGGCAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA
CTACATCCAAATCCTTAGCAGCCCAAACCTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGAGAAAGCG
GAGACAGAGAAAACAAGGGGGGCCAAAGTGGAAGCACATGCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA
CAGAGAACCTGGATGAGCTCCCTAGAGCTGTAGCCACGCTAGTTTTCTACTTAATTTAAAGAAAAGAAATCTCTACC
TATAAAAACATTTGCCCTTCTGTTTTGTATATCTCTATAGTAATTCATAAATGCTTCCCAGTGGAGTTTGTCTAAGG
CACAAGACAATAATCTGAATAAGACAATATGTGATGAATATAAGAAAGGGCAAAAAATTCATTTGAACCAGTTTTT
CCAAGAACAAATCTTGACAAGCAAAGTATAAGAATTATCCTAAAATAGGGGGTTTTACAGTTGTAAATGTTTTTA
TGTTTTTGAGTTTTTGGAAATTTATGTGTCATGTAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGATATAAGGT
GCTTTATTCCCTCGAATGTCCATTAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT
CATTCTCTATTGAGAACCAGCTACCTTTGTGGTAGGGAATAAGAGGTGAGACAAAATTAAGCAACTCCCATTATC
AACAGGAACCTTCTCAGTGAGCCATTCACTCTCGAGAAATGTTATAGAAATTTGGAGAGGTGCAATTATTTCTTTC
TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGGCACTAATGTTTCCCCCAGGATTTCT
ATTGACTAGTCAGGAGTAACAGGTTTCACAGAGAGAAGTTGGTGCCTTAGTTATGTGTTTTTTAGAGTATATACTAA
GCTCTACAGGGACAGAATGCTTAATAAATACCTTAATAAGATATGGGAAAATATTTTAATAAAAACAAGGAAAAACA
TAATGATGTATAATTGCATCCTGATGGGAAGGCATGCAGATGGGAATTTGTTAGAAGACAGAAGGAAAGACGCCAT
AAATCTCTGGCTTTGGGGAAAACCTATATCCCCATGAAAGGAAGAACATCAAAAATAAGGTGAGAGTAATGTAA
TGGAGCTCTTTTCACTAGGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTTTAAAAAAAATCTAGATTATAACA
AACTGCTAGCAAAATCTGAGGAAACATAAATCTTCTGAAGAATCATAGGAAGAGTAGACATTTTATTTTATAACC
AATGATATTTTCAGTATATATTTTCTCTCTTTTTAAAAAATATTTATCATACTCTGTATATTTATTTCTTTTTACTGC
CTTTATTCTCTCCTGTATATTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAAACAATATATAACACACAGA
GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAAATTTTAAC
AACGGAAGGGTTAAATTAACCTTTTGACATCTTCACTCAACCTTTTTCTATTGCTGAGTTAATCTGTTGTAATT
GTAGTATTTGTTTTTTGTAATTTAAACATAAATAAGCCTGCTACATGT

351/615

FIGURE 348

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883

><subunit 1 of 1, 777 aa, 1 stop

><MW: 89651, pI: 7.97, NX(S/T): 3

MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNIPLRLKLTyKDLLLSNSCIPFLG
SSEGLDFQTLTLLDEERGRLLLGAKDHIFLLSLVDLNKNFKKIYWPAAKERVELCKLAGKDANT
ECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFDPPQ
PFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDDHYIRTDISEHYWLNAGAKFIGTFFIPDT
YNPDDDKIYFFFRESSQEGSTSDKTILSRVGRVCKNDVGGQRSLINKWTTFLKARLICSIPGS
DGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSAVCVYSMADIRAVFNPGPYAHKES
ADHRWVQYDGRIPYPRPGTCPSKTYDPLIKSTRDFPDVISFIKRHSVMYKSVYPVAGGPTFK
RINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSSISKEKWNMEEVVLEELQIFKHSS
IILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYGKACADCCCLARDPYCAWDGNACSRYPATSK
RRARRQDVKYGDPIQTCWDIEDSISHETADEKVIFFGIEFNSTFLECIPKSQQATIKWYIQRSG
DEHREELKPDERIIKTEYGLLIRSLQKKDSGMYCYCKAQEHTFIHTIVKLTLNVIENEQMENTQ
RAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRNKGKGPWKHKMQEM
KKKRNRHHRDLDELPRAVAT

Important features of the protein:

Signal peptide:

amino acids 1-36

N-glycosylation sites.

amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.

amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 32-37

20250909 09:50:00

FIGURE 349

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGCAGAGGTCATCCTGGAGCATGCCACCGCGGGGAGCAGA
 CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCACCCAC
 CTCAGCAGTTTCAGCCAGCAGGACTGATCAGGTGTGTGCTCTGGAGTGGGGAGCAGAAGGCGTGGCTGGCAAGA
 GTGGCCTGGAGAAAGAGGTTTACGCGCTTGACCAGCCGAGCTGCCCCGTGACTACAAGATCCAGAACCATGGGCATC
 GGGTGAGGTGGGGGGGCACAGGTGTGATGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
 AGCCATTGAGGGTGTGATGGAGCTACAGAGGGGAGGGAAAGGTATTTTAAGGTAACAGTGTGGCACAATAGTTAA
 GAGCACAGTTTTTTGGAGCTAGACCGACATAGGTTCAAATTCTCTTCTGTTGCTTCCTAGTTCTGTAGCCCCAGGT
 AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCCCTCATCACTAAAGTAGGGCCAATAATAGCACCCACCTCAT
 AGGGAAGATTAAATGACATAATGTATGTGATGCAACTAGCAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAG
 TATTTCCACCCACCCCTGTTCTCTGCCTTCCCAACCAGGTACTGCAACGACTGGAGCAGAGGCGGCAGCAGGCTT
 CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCGAGAGAGCATCCGCCGGGCACAGGTGAGCC
 AGGTGAAGGGGGCTGCCCGGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
 TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGCGGCTCAGTGAGGCTCGGCTGTCCAGAGGGACCTCTCTC
 CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGCAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCCTGCCC
 CCAAGCCCTGGCCACGAGGGCCCTCCCTGCCCTGCACACGTGGTATTTTCGCTATCAGGCAGGGCGTGAGGATG
 AGCTGACAATCACGAGGGTGAAGTGGCTGGAGGTATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA
 ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC
 AAGACAGTGACAATCCCTGCGGGGCAGAGCCACAGCATTCCTGGCACAGGCCCTGTACAGCTACACCGGACAGA
 GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCGGGCCCAAGATGGAGTAGATGACG
 GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCTCCCTGCTGGTGGAGAGCTGCTTGGCCCCC
 CAGGGCCACCTGAACCTCTCTGACCCTGAACAGATGCTGCCGTCCCTTCTCTCTCCAGCTTCTCCCCACCTGCAC
 CTACCTCTGTGTTGGATGGGGCCCCCTGCACCTGTCTGCTGGGGACAAAGCCCTGGACTTCCCTGGGTTCCTGG
 ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACCACCTCCCCCGCCGGCTAAAGCCCCGGATCCTGGCCACC
 CAGATCCCCCTCACCTGAAGGCCAGGGAAGCCTTGACCCCCAGTGATGCTGCTGTCCCTATCTTCAAGCTGTGAGA
 CCACACCATCAATGATCCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCTCACCTCCAAGGGT
 GGAACTTGCCCCCTTCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTCCATTGTTCTATCATCTCTAGGACC
 GGAACCTACTACCTTCTCTTCTGTGATGACCTATCTAGGGTGGTGAATGCCTGAAATCTCTGGGGCTGGAAACC
 ATCCATCAAGGTCTCTAGTAGTTCTGGCCCCCTCTTTCCCCACCTGGCTCCATGACCCACCCACTCTGGATG
 CCAGGGTCACTGGGGTTGGGCTGGGGAGAGGAACAGGCCCTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
 CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAGGGCACGAAGGCCAGGCCAGGGCCTGGGCCCTTTGTG
 CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGGCCACCAGGGCAGGGGCTGCTCCAGCTGACGAC
 GCTCTATCATATGGAGCGAGGTGTTGGGGAAGGCCGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAGAC
 TGAGGGGCTGTGACCTCTCCTGAGGCCCCACAGCCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC
 AGCTGGGGGGCAGTGCTGTCCAGTGGAGGGGAGGGCTTTACGCCCCACCCACCCCTGGCCCTGCCAGCTGGTAG
 TCCATCAGCACAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTTGCTTCCCTGTTCAAGCTGTGTCCAGC
 TTTTCCCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAACCAAGGGATTTATAGCAAAGGCTAAGCCTGC
 AGTTTACTCTGGGGGTTTACGGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA
 TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG
 TCAACAATGAGAGACCAGGAGTAGGTCCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCAGCCAGTGC
 AGTCCCGGCTGTGTTTTCTACCTGGTGATCAGAAGTGTCTGGTTTTGCTTGGCTGCCCATTTGCCCTCTTGAGTGG
 GCAGCCCTGGGCTTGGGGCCCCCTCCCTCCGGCCCTCAGTGTTGGCTCTGCAGAAGCTCTGGGGTTCCTTCAAGTG
 CACGAGGGGTTAGGCTGCTGTCCCTGAGTCCCTCCATTCTGTACTGGGGGGCTGGCTAGGACCTGGGGCTGTGGCC
 TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCAGACCCCTGACCACCCCTG
 GGTCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT
 TCTCAATGTGTGTCACCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACCTCAGAGGCTGCTTG
 GCCCTCCCCTCTGACCAGGGACATCCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCTTCTC
 AGATTGTGGGGCACATTGTGTAGCCTGACTTCTGCTGGAGCTCCAGTCCAGGAGGAAAGAGCCAAGGCCACTT
 TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCTTTCCCTGGAGCACCTGCCCCACCTGCCCA
 CAGAGAACACAGTGGTCTCCCCTGTCCGGGGGCGGCTTTTCTCTCCTTGGAGCGTCCCTGACGGACAAGTGGAG
 GCCTCTTGCTGCGGCTGCAATGGATGCAAGGGGCTGAGAGCCCGAGGTGCAGTGTGTGATGATGGGAGGGGGCTC
 CGTCTGCAAGGTGGAGGTGGCATCCACACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT
 TCATGTTTTGTTTTCTTACGTTCTTTTACGATGCTCCTTAAACCCCAAGCCCCAATTTCCCCAAGCCCCATTT
 TTTCTTGTCTTTATCTAATAAACTCAATATTAAG

353/615

FIGURE 350

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0
MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLQRRQQASEREAPSIEQRLQEVRESIRRAQ
VSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELSDF
EECEETGELFEPPAPQALATRALPCPAHVVFQAGREDELTITEGEWLEVIEEGDADEWVKA
RNQHGEVGVFPERYLNFPDLSLPSSQSDNPCGAEPFLAQALYSYTGQSAEELSFPEGAL
IRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPSFSPPA
PTSVLDGPPAPVLPDGDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT
```

seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401

FIGURE 351A

CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCACAGAGTCCCTACAGAGGGAGAGGCCAGAGAAGCTGCAGA
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCTTTAACGGTCTTCAGCCCTTCTGAAAACTTTGCC
TCTGACCTTGGCAGGAGTCCAAGCCCCAGGCTACAGAGAGGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT
GCCCTAGACGGCCTCAGTCCCTCCAGCTGCAGTACCAGTGCC**ATG**TCCAGACAGGCTCGCATCCCGGGAGGGG
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCCTGCCCTCCTGCTCCCATTTGTGCCGCTCTCCTGGCTGGTGTG
GCTGCTTCTGCTACTGCTGGCCTCTCTCCTGCCCTCAGCCCGGCTGGCCAGCCCCCTCCCCGGGAGGAGGAT
CGTGTTCAGAGAAGCTCAACGGCAGCGTCTTGCTGGCTCGGGCGCCCCCTGCCAGGCTGTTGTGCCGCTTGCA
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTGAGGGGGCTGACAGTGCAGTA
CCTGGGGCAGGCGCCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTTAGGCGTGTACAAATATCGGGGGCTGAACTCCA
CCTCCAGCCCCCTGGAGGGAGGCACCCCTAACTCTGCTGGGGGACCTGGGGCTCACATCCTACGCCGGAAGAGTCC
TGCCAGCGGTCAAGTCCCCTGTCACAGTCAAGGCTCCTCTTGGGAAGCCCCAGCCCCAGACCCGAAGAGCCAA
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCTTCCAGCGTGC
GGGGCTAAAGCGCTACCTGCTAACAGTGATGGCAGCAGCAGCCAAGGCCTTCAAGCACCCAAGCATCCGCAATCC
TGTGAGCTTGGTGGTGAAGTGGCTAGTGATCCTGGGGTCAAGCGAGGAGGGGGCCCCAAGTGGGGGGCAGTGTGC
CCAGACCCCTGCGCAGCTTCTGTGCTGGCAGCGGGGCTCAACACCCCTGAGGACTCGGGCCCTGACCACTTTGA
CACAGCCATTCTGTTTACCCGTGAGGACCTGTGTGGAGTCTCCACTTGCAGACAGCTGGGTATGGCTGATGTGG
CACCCTCTGTGACCCGGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTGCTCA
TGAATGGGTGATGTCTTCAACATGCTCCATGACAACCTCCAAGCCATGCATCAGTTTGAATGGGCCTTTGAGCAC
CTCTCGCCATGTGATGGCCCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGGCCGCTT
CATCACTGACTTCTTGGACAATGGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATTGCATCTGCCCTGT
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCCAGTGCCAGCTGACCTTCGGGGCCGACTCACGCCATTGTCC
ACAGCTGCCGCCGCCCTGTGCTGCCCTCTGGTGCTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAACA
CTCGCCCTGGGCCGATGGCACACCCCTGCGGGCCCGCACAGGCCTGCATGGGTGGTCTGCTCCCATGGACCA
CTCCAGGACTTCAATATTTCCACAGGCTGGTGGCTGGGGTCTTGGGGACCATGGGGTGAATGCTGCTCTCGGACCTG
TGGGGGTGGTGTCCAGTTCTCCTCCCGAGACTGCACGAGGCCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA
GTGTGCTGCCCTACAACCACCGCACCGACCTCTTCAAGAGCTTCCAGGGGCCATGGACTGGGTTCCTCGCTACAC
AGGCGTGGCCCCCAGGACCAGTGCAAACTCACCTGCCAGGCCGGGCACTGGGCTACTACTATGTGCTGGAGCC
ACGGGTGGTAGATGGGACCCCTGTTCCTCCGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGCTGG
CTGTGATCGCATGATTGGCTCCAAGAAGAAGTTTGACAAGTGATGGTGTGCGGAGGGGACGGTTCTGGTTGCAG
CAAGCAGTCAGGCTCCTTCAGGAAATTCAGGTACGGATAACAATGTGGTCACTATCCCCGCGGGGGCCACCCA
CATTCCTGTCCGGCAGCAGGGAAACCCCTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCCTA
TGCCCTCAATGGTGAATACACGCTGATGCCCTCCCCACAGATGTGGTACTGCCCTGGGGCAGTCAGCTTGGCTA
CAGCGGGGCCACTGCAGCCTCAGAGACACTGTGAGGCCATGGGGCACTGGCCAGCCCTTTGACACTGCAAGTCTT
AGTGGCTGGCAACCCCCAGGACACACGCCCTCCGATACAGCTTCTTCTGCCCCGGCCGACCCCTTCAACGCCACG
CCCCACTCCCCAGGACTGGCTGCACCGAAGAGACAGATTCTGGAGATCCTTCGGCGGCGCCCCCTGGGCGGGCAG
GAAAT**TA**ACCTCACTATCCCGCTGCCCTTTCTGGGCAACCGGGGCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT
CTGTTGCTGCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCCCTCCTCTGCCCCAAT
GCGCAGGCTGGCCCTGCCCTGGTTTCTGCCCCGGGAGGCACTGATGGGTAGTGGATGGAAGGGGCTGACAGAC
AGCCCTCCATCTAAACTGCCCCCTCTGCCCTGCGGTCACAGGAGGGAGGGGGAAGGCAGGGAGGGCCTGGGCCC
CAGTTGTATTTATTTAGTATTTATTTACTTTTATTTAGCACAGGGAAGGGGACAAGGACTAGGGTCTTGGGGAA
CCTGACCCCTGACCCCTCATAGCCCTCACCCTGGGGCTAGGAAATCCAGGGTGGTGGTATAGGTATAAGTGGTG
TGTGTATGCGTGTGTGTGTGTGTGTGAAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCTCT
TTCCTGAATTTTATTTTGGGAAAAGAAAAGTCAAGGGTAGGGTGGGCCTTCAGGGAGTGAGGGATTATCTTTT
TTTTTTTTTTCTTTCTTTCTTTCTTTTTTTTTTTTTTGTAGACAGAATCTCGCTCTGTGCGCCAGGCTGGAGTGCAATG
GCACAATCTCGGCTCACTGCATCCTCCGCTCCCGGTTCAAGTGATTCTCATGCCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCTCCTGCCACCACGCCAGCTAATTTTGTGTTTGTGTTTGGAGACAGAGTCTCGATTCTGCTC
ACCAGGCTGGAATGATTTTCACTGACCTGCAACTTCCGACCTTGGGTTCCAGCAATTCTCCTGCCTCAGCCTCC
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCGGCTAATTTTTGTATTTTGTAGTAGAGACGGGGTTTTCAC
CATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTAGGTGATCCACTCGCCTTCATCTCCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGTGCTGGCCACGCCCACTAATTTTTGTATTTTGTAGTAGAGACAGGGTTTACCATGT
TGGCCAGGCTGCTCTTGAACCTGACCTCAGGTAATCGACCTGCCCTCGGCCTCCCAAAGTGCTGGGATTACAGG

356/615

FIGURE 352

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492
><subunit 1 of 1, 837 aa, 1 stop
><MW: 90167, pI: 8.39, NX(S/T): 1
MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLLASLLPSARLASPLPREEEIVF
PEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEPGT
YLTGTINGDPESVASLHWDGGALLGVLYRGAELHLQPLEGGTPNSAGGPGAHILRRKSPASG
QGPMCNVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAAAKA
FKHPSIRNPVSLVVTSLVILGSGEEGPQVGPSAAQTLRSFCAWQRGLNTPEDSGPDHFDTAIL
FTRQDLGCVSTCDTLGMADVGTVCDFPARSCAIVEDDGLQSAFTAHELGHVFNMLHDNSKPCI
SLNGPLSTSRHVMAVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGK
DYDADRQCQLTFGPDSRHCPQLPPCAALWCSGHLNGHAMCQTKHSPWADGTPCGPAQACMGG
RCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRPVPRNGGKYCEGRRTRFR
SCNTEDCPTGSALTFFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCKLTCQARALGY
YYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGDGS GCSKQSGSFR
KFRYGYNNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTLMPSPTDVVLPGA
VSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPTPSTPRPTPDWLH
RRAQILEILRRRPWAGRK

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 358-367

20050326-101502

357/615

FIGURE 353

GCGGAAC TGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCCC
GGCTGGCCTAGGCAGGCAGCCGCACCATGGCCAGCACGGCCGTGCAGCTTCTGGGCTTCCTGC
TCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCAACCATCCTGCCGCACTGGCGGAGGACAG
CGCACGTGGGCACCAACATCCTCACGGCCGTGTCCTACCTGAAAGGGCTCTGGATGGAGTGTG
TGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCCCCAAG
ACCTCCAGGCTGCCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCTGCGCCT
GCGCCGT CATCGGGATGAAGTGACGCGCTGCGCCAAGGGCACACCCGCCAAGACCACCTTTG
CCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGTCTCCTGGA
CCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCAGCGGCATGAAGTTTGAGA
TTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGTGGCACCTGC
TTTGCCTGTCCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCCAGGGCCACCA
CGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAGACAATCGGGCCC
CCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTGTGAGTCCCCACAG
CCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTCAATGGAGGCAGGGG
TTCCAGCACAAAGTTTACTTCTGGGCAATTTTGTATCCAAGGAAATAATGTGAATGCGAGGA
AATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAAAGCTCTCTATACCAA
GACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATATTTATGTGGGTGATTTGA
TAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGTTGTTTTGTGATCCAGGAA
TAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

358/615

FIGURE 354

MASTAVQLLGFLLSFLGMVGTLITTLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIYQ
CQIYRSLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAGTTPAKTTFAILGGTLFI
LAGLLCMVAVSWTTNDVVQNFYNPLLPSGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQDEA
PYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GAGCTCCCTCAGGAGCGCGTTAGCTTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGC
AGGCGGCAGGGCGGGCGGCCAGGATC**ATGT**CCACCACCACATGCCAAGTGGTGGCGTTCCTCC
TGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGACCCAGG
ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA
GGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGCCATGC
TGCAGGCAGTGCAGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTAT
CCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGA
CACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGT
TTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTG
GGATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGC GGCTCTGTTTCGTGGGCTGGGTG
CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAG
AAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTG
GAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATATACGATGGAG
GTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTATGT**GTAAT**GCTCTA
AGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGATCCCA
TCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTG
GAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG
TTATTTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTTTAAATATAA
CTTTCTACTCTGATGAGAGAATGTGGTTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAG
ACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG
AAAACTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTC
CCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAAAGATATT
TGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAAGTTACCAAACCA
AAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGC
AACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTTCTGTGCGGGGTGAGAAA
TTGTCCCTAGATGAATGAGAAAATTATTTTTTTTTTAATTTAAGTCCTAAATATAGTTAAAATAA
ATAATGTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG
AAGGAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAA
GTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGG
ATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAA
ATACAGAGAGAAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCCGGGAG
GCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACC
ACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGA
ACACAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTTAA

360/615

FIGURE 356

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734

><subunit 1 of 1, 261 aa, 1 stop

><MW: 27856, pI: 8.50, NX(S/T): 1

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGWLWRSCVRQSSGFTEC
RPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFI
VSGLCIAIGVSVFANMLVTNFWMSTANMYTGMGMVQTVQTRYTFGAALFVGWVAGGLTLIGG
VMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQ
SYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 357

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGG
AGTCCAGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGG
CTGTTTCTTGGTGGTGTGGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGA
GTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAAACCTTCTGGGAAGGACTGTGGATG
AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTT
TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCT
TTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAG
GCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGTCTCATCCCT
GTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAA
AAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTCTGATTGTTGGA
GGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA
CCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCCGAAAGAAGTCACCGAGCGTCTACTCC
AGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTTAACTTTACTATAAAGCCATGCAATGACA
AAAATCTATATTACTTTCTCAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAACTGC
CTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTTACGAGAA
TGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGG
TTCAAGCATCTACTCTTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTATTTT
ACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCA
CATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAGTCCATTACACTGAATAAAT
AGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTGAAGAAGGTTACTATTAATTG
TTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGAAGATTAAAATGAAGGCTTTAA
TCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGTTTTTTTAGCCTAGGAGTTAGAA
ATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTTTCTTGTGTATTAAATTAACATT
TTTAAAACGCAGATATTTTGTCAAGGGGCTTTGTCATTCAAACCTGCTTTTCCAGGGCTATACTC
AGAAGAAAGATAAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTTAGGAAAGTGAAAATATTTTT
GTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGATATATTT
TAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATAT
GTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTC
ACCTGCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTTT
CTTCTTGTACCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATT
CCAAATTTGATGAAACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAAT
GAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTGTACATTTTTTCTAATT

362/615

FIGURE 358

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

><subunit 1 of 1, 225 aa, 1 stop

><MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQ
CKIYDSL LALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIF
IITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCN
EKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

363/615

FIGURE 359

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCTCGCCCTCCTGCTCGCGCCCC
GCCGCCATGGCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGCTC
CTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAAATAAACTCAAGCTGATGCTTCAAAA
CGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGAATTC
CTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGCAGCAG
TGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATCACCTAT
TGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCACTATGAT
GAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCGTCAACTAC
GATGACTACTAACCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTCTCTTCATGT
ATCTCCTAATGCCTTACACTACTTGGT'TTCTGATTTGCTCTATTTTCAGCAGATCTTTTCTACC
TACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTTTGATATTTTCAT
GGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTTAAAAAGA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

364/615

FIGURE 360

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLLMLQKREAPVPTKTKVAVDENKAKEFLG

SLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRNGHEYDYYQRHYDED

SAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 361

[illegible]

366/615

FIGURE 362

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCLAFVESKFNISKIN
ENADGSFDYGLFQINSHYWCNDYKSYSENLCBVDCQDLLNPNLLAGIHCAKRIVSGARGMNNW
VEWRLHCSGRPLSYWLTGCRLR

Signal peptide:

amino acids 1-18

2025-10-10 10:10:10

FIGURE 363

TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGACTGAGAACCCAC
 CAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAGCA
 GAATGGAATCATTATTTTTTTTCCCAAGGAGAAAAACCGGGGTAAAGGGAGGGAAGCAATTCAAT
 TTGAAGTCCCTGTGAATGGGCTTTTTCAGAAGGCAATTAAAGAAATCCACTCAGAGAGGACTTGG
 GGTGAAACTTGGGTCTGTGGTTTTCTGATTGTAAGTGAAGCAGGTCTTGACACACGCTGTTG
 GCAAATGTCAGGACCAGGTTAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCAACCCAT
 GTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGACCTGTTGC
 ATTTGGCAAGTTCTAGCAAC**ATG**CTCCTAAGGAAGCGATACAGGCACAGACCATGCAGACTCC
 AGTTCCTCCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGTTGCACCCTC
 CCCACCACACCCTGCACCAGACTGTACAGCCCAAGCCAGCAAGCACAGCCCTGAAGCCAGGT
 ACCGCCTGGACTTTTGGGGAATCCCAGGATTGGGTACTGGAAGCTGAGGATGAGGGTGAAGAGT
 ACAGCCCTCTGGAGGGCCTGCCACCCCTTTATCTCACTGCGGGAGGATCAGCTGCTGGTGGCCG
 TGGCCTTACCCCAGGCCAGAAGGAACCAGAGCCAGGGCAGGAGAGGTGGGAGCTACCGCCTCA
 TCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGGGGCTGATGAGGACG
 GGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTTCAGCCTGGACCCACGTGGCCTCCAGGAGG
 CACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCCAGGTGCGGCACCCACTGTGTCTGC
 AGCAGCACCCCTCAGGACAGCCTGCCCCACAGCCAGCGTCATCCTCTGTTTCCATGATGAGGCCT
 GGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGCCAGGGCCTTCCCTGAAGG
 AGATCATCCTCGTGAGCAGCCTCAGCCAGCAAGGACAACTCAAGTCTGCTCTCAGCGAATATG
 TGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAGGCTGGGTGCCATCAGGGCCC
 GGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCTTCATGGATGCCCACTGCGAGT
 GCCACCCAGGCTGGCTGGAGCCCCCTCCTCAGCAGAAATAGCTGGTGACAGGAGCCGAGTGGTAT
 CTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTATTACCCCTCAAAGGACCTGCAGC
 GTGGGGTGTGGACTGGAAGCTGGATTTCCACTGGGAACCTTTGCCAGAGCATGTGAGGAAGG
 CCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTGGTGCCCGAGAGGTGGTGGCCATGG
 ACAGACATTACTTCCAAAACACTGGAGCGTATGACTCTCTTATGTGCTGCGAGGTGGTGAAG
 ACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTGGCTCTGTTGAAATCCTTCCCTGCTCTC
 GGGTAGGACACATCTACCAAATCAGGATTCCCATTTCCCCCTCGACCAGGAGGCCACCCTGA
 GGAACAGGGTTTCGATTGCTGAGACCTGGCTGGGGTCATTCAAAGAAACCTTCTACAAGCATA
 GCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGAAGCCAGACTGCATGGAACGCTTGCAGCTGC
 AAAGGAGACTGGGTTGTGCGACATTCCACTGGTTTCTGGCTAATGTCTACCCTGAGCTGTACC
 CATCTGAACCCAGGCCAGTTTCTCTGGAAGCTCCACAACACTGGACTTGGGCTCTGTGCAG
 ACTGCCAGGCAGAAGGGGACATCCTGGGCTGTCCCATGGTGTGGCTCCTTGCAGTGACAGCC
 GGCAGCAACAGTACCTGCAGCACACCAGCAGGAAGGAGATTCACTTTGGCAGCCCACAGCACC
 TGTGCTTTGCTGTCAGGCAGGAGCAGGTGATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCA
 TCCACCAGCAGCACTGGGACTTCCAGGAGAATGGGATGATTGTCCACATTCTTTCTGGGAAAT
 GCATGGAAGCTGTGGTGAAGAAAACAATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAG
 CCCGCCAGCAGTGGCGATTGACCAGATAAATGCTGTGGATGAACGAT**TGA**ATGTCAATGTGAG
 AAGGAAAAGAGAATTTTGGCCATCAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATT
 TCATGAAGCTGATCCTTTTGTGTGTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAA
 GAATATAGGAAGTTTCTCCTTTTTCACACCTTATTTCAATTGACTGCTGGCTGCTTA

368/615

FIGURE 364

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFGE
SQDWVLEAEDEGEEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPRRQ
DKEAPKRDWGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHPQDS
LPTASVILCFHDEAWSTLLRRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVARLEGV
KLLRSNKRLGAIRARMLGATRATGDVLFVMDAHCECHPGWLEPLLSRIAGDRSRVVSFVIDVI
DWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMDRHYFQN
TGAYDSLMSLRGGENLELSFKAWLCGGSVEILPCSRVGHYQNDSSHSPLDQEATLRNRVRIA
ETWLGSFKETFYKHSPEAFSLSKAEKPCMERLQLQRRRLGCRTFHWFLANVYPELYPSEPRPS
FSGKLHNTGLGLCADCQAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSPQHLCFAVRQ
EQVILQNCTEEGLAIHQHWFQENGMIVHILSGKCMEAVVQENNKDLYLRPCDGKARQQWRF
DQINAVDER

Signal peptide:

amino acids 1-28

20510-99525001

369/615

FIGURE 365

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGG
AGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGC
AGCCGGGAGCC**ATG**CGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGC
TGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAA
AGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAG
CAGGAGTGCCTGGTTCGAGACGGGAGCCCTGGGGCCAATGTTATTCGGGTACACCTGGGATCC
CAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCT
GGACACCCAACTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAA
TTGCGGAGTGACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCAGTGGCT
CACTTCGGCTAAATATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTG
AATGTTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGA
ATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTG
GATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTA
CTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAA**TAA**ATGCTTTAATTTT
CATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCATTAAATGACATTTTAAATAAG
TTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGACCAAAGTGTGATTTCA
CACTGTTTTTAAATCTAGCATTATTCAATTTTGCTTCAATCAAAGTGTTTCAATATTTTTTT
TAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTATAATTTGGAATATTGTTG
TGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTAAAAAATATAAAAGCTACCAATCTTTG
TACAATTTGTAAATGTAAAGAATTTTTTTTTATATCTGTAAATAAAAATTATTTCCAACA

2025-10-23 10:43:00

370/615

FIGURE 366

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVP
GRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAEC
TFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPPEMNSTI
NIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRITIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 367

[illegible]

372/615

FIGURE 368

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398

><subunit 1 of 1, 121 aa, 1 stop

><MW: 12073, pI: 4.11, NX(S/T): 0

MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTP

PAPTVPAGPEDSTAQERLDQGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 91-110

Glycosaminoglycan attachment site.

amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 116-119

N-myristoylation site.

amino acids 91-96

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGCCGTTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGCGG
GGGCCGGGACGGGCATGGCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCACG
GCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGAACT
TCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAAGGGCGCTGCTCACCGACTGGAGCGACG
ACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGACCAAG
TGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCCCCGGGT
ATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAACGCCATCA
TCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGGACCCAGCC
TAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCCTAAGTAGCCCCCAGAGGCG
CTGGGAGTGTTGCCACCGCCCTCCCCTGAAGTTTGCTCCATCTCAGCTGGGGGTCAACCTGG
GGACCCCTTCCCTCCGGGCCATGGACACACATAACATGAAAACCAGGCCGCATCGACTGTCAGC
ACCGCTGTGTCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAGACTCGCACGTCTG
CCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGTGCCAGGGCCCTAC
TGTCCCTGGGGTCCCAGGCTCTCCTTGAGGGGGCTCCCCGCCTTCCACCTGGCTGTTCATCGG
GTAGGGCGGGGCCGTGGGTTCAAGGGCGCACCACTTCCAAGCCTGTGTCCACAGGTCTTCGG
CGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAAATAACTGGCACAAAGTAAGTCC
CCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTGGGTGAGTATGTGTGGGG
CACAGGCTGGCTCCCTCAGCTCCACGTCTTAGAGGGGCTCCCGAGGAGGTGGAACCTCAACC
CAGCTCTGCGCAGGAGGCGGCTGCAGTCTTTTTCTCCCTCAAAGGTCTCCGACCCCTCAGCTGG
AGGCGGGCATCTTTTCTAAAGGTCCCCATAGGGTCTGGTTCCACCCCATCCAGGTCTGTGG
TCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCCATGGAGGGGCTGACTGCCCCA
CATTGCCTTTTTCAGACAGGACACGAGCATGAGGTAAGGCCGCCCTGACCTGGACTTCAGGGGGGA
GGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCTTAGATCAGTGGGGGCACTGCAGGT
GGGGCTCTCCCTATACTGGGACACCTGCTGGATGTCACCTCTGCAACCACACCCATGTGGTG
GTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGGCCTGGGACACACAGAGCCACCCCGG
CCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGAAGGGGTGCTCGTAAGCCAACACCAGC
GTGCCGCGGCCTGCACACCCCTTCGGACATCCCAGGCACGAGGGTGTCTGTGGATGTGGCCACAC
ATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCTGGGGCCCCCAGGGAGGGAGGCAGGGGGT
GGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCCCCGCAGCCTGGTATCGCCAGCCTTAAGGT
GTCTGGAGCCCCCACACTTGGCCAACCTGACCTTGGGAAGATGCTGTCTGAGTGTCTCAAGCAGC
ACTGACAGCAGCTGGGCCTGCCCCAGGGCAACGTGGGGGGCGGAGACTCAGCTGGACAGCCCT
GCCTGTCACTCTGGAGCTGGGCTGTCTGCTGCCTCAGGACCCCCCTCTCCGACCCCGGACAGAGC
TGAGCTGGCCAGGGCCAGGAGGGCGGGAGGGAGGGAATGGGGGTGGGCTGTGCGCAGCATCAG
CGCCTGGGCGAGGTCCGCAGAGCTGCGGGATGTGATTAAAGTCCCTGATGTTTTCTC

374/615

FIGURE 370

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCLVCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMKE

LHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIERHL

APGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

20250304 10:50:00

375/615

FIGURE 371

GCCGGCTGTGCAGAGACGCCATGTACCGGCTCCTGTCAGCAGTGA CTGCCCGGGCTGCCGCCC
CCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCCTC
TCGGCCACGGCTGGGTGCGGGGCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGCTGG
CAGGTGGGCTGAGGGGCGCGGCCCGGCGCAGTCCCCCGCGGCCCCCGACCCCTGAGGCGTCGC
CTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGCGCCGC
CCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGATCAAGG
ATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTCTGGTCAG
AAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGTTATGCGAA
TTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAGCAGGGAAAC
TGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCAGAGAAAAGAATATGAAGGTG
AAAAGGTTTCTGTCAACAAGATTACTGATTTCCCATTTAAGTGGAAATTCGTCATTATGAAA
AGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGATGAAAGAGAATG
TTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTTACTAAATTTAAAA
CAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAAAAAGAATGATTTTG
AACAAGGCGAATTATATTTGAGAGAAAAGTTTGAATTTCAATTGAATCCCTAAGATTATTTA
AAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCAACTTTTGGCTATACCC
TACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGACTATATGCAGAAAATAT
TCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGCCAGTGATTTACAATAGAG
CAAGGTAAATGAATACCTTCTGCTGTGTCTAGCTATATCGCATCTTAACACTATTTTATTAAT
TAAAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGCCACATTTTGGGAGCTTTTCT
ACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACATGTTTATAAAGTAAAAAAA

2025-10-27 10:43:00

376/615

FIGURE 372

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41221, pI: 8.54, NX(S/T): 0

MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGGLGLGLGLALGVKLAGGLRGA
APAQSPAAPDPEASPLAEPPEQSLAPWSPQTPAPPSCRCFARAIESSRDLLHRIKDEVGAPG
IVVGVSVVDGKEVWSEGLGYADVENRVPCPETVMRIASISKSLTMVALAKLWEAGKLDLDIPV
QHYVPEFPEKEYEGEKVSVTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAFEQEK
EGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKNDPLFF
KPGSQFLYSTFGYTLAAIVERASGCKYLDYMQKIFHDLDMMLTTVQEENEPVIYNRAR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 39-60

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

377/615

FIGURE 373

GACTACGGGGAGAGAGAGAGAGACCAGGACAGCTGCTGAGACCTCTAAGAAGTCCAGATACTAA
GAGCAAAGATGTTTTCAAACCTGGGGGCCTCATTGTCTTCTACGGGCTGTTAGCCCAGACCATGG
CCCAGTTTGGAGGCCTGCCCCGTGCCCCTGGACCAGACCCCTGCCCTTGAATGTGAATCCAGCCCTG
CCCTTGAGTCCCACAGGTCTTGACAGGAAGCTTGACAAATGCCCTCAGCAATGGCCTGCTGTCT
GGGGGCCTGTTGGGCATTCTGGAAAACCTTCCGCTCCTGGACATCCTGAAGCCTGGAGGAGGT
ACTTCTGGTGGCCTCCTTGGGGGACTGCTTGGAAGTGACGTCAGTGATTCTGGCCTGAAC
AACATCATTGACATAAAGGTCACCTGACCCCCAGCTGCTGGAACCTTGGCCTTGTGCAGAGCCCT
GATGGCCACCGTCTCTATGTCAACATCCCTCTCGGCATAAAGCTCCAAGTGAATACGCCCCTG
GTCGGTGCAAGTCTGTTGAGGCTGGCTGTGAAGCTGGACATCACTGCAGAAATCTTAGCTGTG
AGAGATAAGCAGGAGAGGATCCACCTGGTCCCTGGTGACTGCACCCATTCCCCTGGAAGCCTG
CAAATTTCTCTGCTTGATGGACTTGGCCCCCTCCCCATTCAAGGTCTTCTGGACAGCCTCACA
GGGATCTTGAATAAAGTCCTGCCTGAGTTGGTTCAGGGCAACGTGTGCCCTCTGGTCAATGAG
GTTCTCAGAGGCTTGGACATCACCCCTGGTGCATGACATTGTTAACATGCTGATCCACGGACTA
CAGTTTGTCAAGGTCTTAAGCCTTCCAGGAAGGGGCTGGCCTCTGCTGAGCTGCTTCCCAG
TGCTCACAGATGGCTGGCCCATGTGCTGGAAGATGACACAGTTGCCTTCTCTCCGAGGAACCT
GCCCCCTCTCCTTTCCCACCAGGCGTGTGTAACATCCCATGTGCCTCACCTAATAAAATGGCT
CTTCTTATGCA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

378/615

FIGURE 374

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76533

><subunit 1 of 1, 256 aa, 1 stop

><MW: 26713, pI: 5.62, NX(S/T): 0

MFQTGGLIVFYGLLAQTMAQFGGLPVPLDQTLPLNVNPALPLSPTGLAGSLTNALSNGLL
SGLLGILENLPLLDILKPGGGTSGLLGGLLGKVTSVIPGLNNIIDIKVTDPQLLELGL
VQSPDGHRLYVTIPLGIKLQVNTPLVGASLLRLAVKLDITAEILAVRDKQERIHVLGDC
THSPGSLQISLLDGLGPLPIQGLLDSLTGILNKVLPPELVQGNVCPLVNEVLRGLDITLVH
DIVNMLIHGLQFVIKV

Important features of the protein:

Signal peptide:

Amino acids 1-19

Transmembrane domain:

Amino acids 79-97

N-myristoylation sites:

Amino acids 46-52;49-55;58-64;62-68;66-72;80-86;81-87;
82-88;85-91;86-92;89-95;202-208;233-239

FIGURE 375

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTTT
AAAT**ATGT**CAAGATCCAGACTTTTCAGTGTACCTCAGCGATCTCAACGATAGGGATCTTGTG
TTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGTAAA
TTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTTGGTGGC
AGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGAATTGA
TTCTCACAGGCGCACCATGGCAGTTTTTGTCTGTTGGAGACTTGGACTCTATTTATGGGACAGA
AGCAGCTGTGAGTCCAACCTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTATATCCTGT
TCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTAAAAACAAC
CTGATTTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCAAAACATCAA
ATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATAGTCTGAGGAA
GGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATGTTTAAATAGTA
AAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACACAACCAAATTAAT
GCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTTTCCATCACATTTA
GGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTCCTAGCATGGGGTCC
ATAAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACATGTCCAGAACCAGAAC
CAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAAGAAAAGTTTGGAGTTG
AAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCTCAGCAAAAACAAGAGGT
TTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGAAGGAGATTGCTGAAGATA
TAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCCCTTTTACAGTAATGAATGT
GGCTCCATAGTCCATAGTGTCTCTGGAGCCTCAGGGCTTGGCATTATTTGCAGCATCATG
CTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGCAGAAGTAGCAATGAGACATCT
TCAAGTGGCATTTTGGCAGTGGCCATCAGCAGGGGGACAGACAAAACATCCATCACAGATGA
CATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAATAAACATCAATAGATATCTAAAAA

380/615

FIGURE 376

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCCEEDEMCVNYNDQHPNGWYIWILLLLVLVAA
LLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYPVP
APCFGPLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

Downloaded from www.jstor.org

FIGURE 377

CGCGGATCGGACCCAAGCAGGTCTGGCGGCGGCGGCAGGAGAGCGGCCGGGCGTCTCAGCTCCTCG
ACCCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCC**ATG**GCCAGGCCCGGC
ATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGCC
GTGGCCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAAC
ATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGA
TGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC
GGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGC
ACCAGTGGTTTGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAA
GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATG
TCTGGCCACCGAGTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTC
ACTGCGCTGACAGAGGGACTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACG
TGATCTCTCCAGGTGTGGTGGAGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAG
AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTT
ATCTACGTCCTCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAG
CAGGTGACCT**TAGT**GACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCCTG
CCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACC
AGGGGCTAGAAAATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATG
TGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTG
TGCCCTGGGCACCTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTT
GTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG
GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCC
ACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCCAGTCTTGCTTCTTGT
CCCCCTCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC
CCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAAAAAAAAAAAAA

382/615

FIGURE 378

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGT
LIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALS
ICTREAYQSMKERNVDDGHIININMSGHRVLP LSVTHFYSATKYAVTALTEGLRQELREAQT
HIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQ
MRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 379

GAGCGGAGTAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCACCGGCGGTT
TCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCGCTTATTAGC
TCTCGCTGCGTCGCCCGGCTCAGAAGCTCCGTGGCGGCGGCGACCGTGACGAGAAGCCACG
GCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAGTCAGATGCCCTTTTAAACTCCCT
CTTCAAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATACAACCTTGCTGAAGATGAA
GAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTTCAAGTCTTGATTTGTGGCTTACCT
CAAGTTACCATTTTTTCAGTCAAGTCTGTTTGTTTGCTTCTTCAGAAATGTTTTTTTACAATCTC
AAGAAAAAATATGTCCAGAAATTGAGTTTACTGTTGCTTGATTTGGACTCATTTGGGGATT
GATGTTACTGCACTATACTTTTCAACAACCAAGACATCAAAGCAGTGTCAAGTTACGTGAGCA
AATACTAGACTTAAGCAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGAACACAGTGG
TGTCGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAAATTGCTGTCTTCT
GGATGACATTTTGCAACGATTGGTGAAGCTGGGAGAACAAAGTTGACTATATTGTTGTGAATGG
CTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAACCACAAATAAAAG
AACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGCTGCTGCTCCATCCACTG
TGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGACAGAGCAATACTTTAC
AATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTGATGAACTCTAATTCTGTACATA
AAAATTTTAAAGTTATTTGTTTGCTTTCAGGCAAGTCTGTTCAATGCTGTACTATGTCTTAA
AGAGAATTTGGTAACTTGTTGATGTGGTAAGCAGATAGGTGAGTTTTGTATAAATCTTTTGT
GTTTGAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATTGATTTCTATAACACATTTAT
TTAAGTATATAACACGTTTTTTTGACAAGTGAAGAATGTTTAATCATTCTGTGATTTGTTCTC
AATAGATGTAAGTGTAGACTACGGCTATTTGAAAAAATGTGCTTATTGTACTATATTTTGT
ATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAATAATGTTTTGAAATCATGACCC
AAAGAATGTATTGATTTGCACTATCCTTCAGAATAACTGAAGGTTAATTATTGTATATTTTTA
AAAATTACACTTTATAAGAGTATAATCTTGAAATGGGTAGCAGCCACTGTCCATTACCTATCGT
AAACATTGGGGCAATTTAATAACAGCATTAAAATAGTTGTAAACTCTAATCTTATACTTATTG
AAGAATAAAAGATATTTTTATGATGAGAGTAACAATAAAGTATTCATGATTTTTTCACATACAT
GAATGTTGATTTAAAAGTTTAATCCTTTGAGTGTCTATGCTATCAGGAAAGCACATTATTTCC
ATATTTGGGTAAATTTTGCTTTTTATTATATTGGTCTAGGAGGAAGGGACTTTGGAGAATGGAA
CTCTTGAGGACTTTAGCCAGGTGTATATAATAAAGGTACTTTTGTGCTGCATTAAATTGCTTG
GAAAGTGTTAACATTATATTATATAAGAGTATCCTTTATGAAATTTTGAATTTGTATAACAGA
TGCATTAGATATTGATTTTATATAATGGCCACTTAAAATAAGAACATTTAAAATATAAACTAT
GAAGATTGACTATCTTTTCAGGAAAAAAGCTGTATATAGCACAGGGAACCCTAATCTTGGGTA
ATTCTAGTATAAAACAAATTATACTTTTATTTAAATTTCCCTTGTAGCAAATCTAATTGCCAC
ATGGTGCCCTATATTTCATAGTATTTATTCTCTATAGTAACTGCTTAAGTGCAGCTAGCTTCT
AGATTTAGACTATATAGAATTTAGATATTGTATTGTTGCTCATTATAATATGCTACCACATGT
AGCAATAATTACAATATTTTATTAATAAATAAATATGTGAAATATTGTTTCATGAAAGACAGATT
TCCAAATCTCTCTTCTCTTCTCTGTACTGTCTACCTTTATGTGAAGAAATTAATTATATGCCA
TTGCCAGGT

FIGURE 380

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648
```

```
><subunit 1 of 1, 140 aa, 1 stop
```

><MW: 15668, pI: 10.14, NX(S/T): 5

MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE

NKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNLVP

VT'NKRTNVSGSIR

Important features of the protein:

Signal peptide:

amino acids 1-26

FIGURE 381

AACTTCTAC**ATG**GGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACC
ATCATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGTGCAGAGTCTCAGTTGCCCGG
GAGCACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTT
TCATGCCAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCATTTAAGAGGGTT
TTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAG
CCACCTGTGGCCGTTGAGTGCTTGAAA**TGA**GGAAGTGAAGAAAATTAATTTCTCATGTATTTTT
CTCATTTATTTATTAATTTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGG
ATACATGTATACAATATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACAT
TTATTTTTTTATTCCTTTTTTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCC
ATCTCAGCTTACTGCAACCTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAA
GTAGCTGGGACTACAGGCATGCACCACAATGCCCAACTAATTTTTGTATTTTTTAGTAGAGACG
GGGTTTTGCCATGTTGCCCAGGCTGGCCTTGAAGTCCCTGGCCTCAAACAATCCACTTGCCTCG
GCCTCCCAAAGTGTTATGATTACAGGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTT
CTTTGTGTTGGGAACTTTGAAATTATACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTG
CTATGGAACACTGGGACTTCTTCCCTCTATCTAACTGTATATTTGTACCAGTTAACCAACCGT
ACTTCATCCCCACTCCTCTCTATCCTTCCCAACCTCTGATCACCTCATTCTACTCTCTACCTC
CATGAGATCCACTTTTTTTAGCTCCACATGTGAGTAAGAAAATGCAATATTTGTCTTTCTGTG
CCTGGCTTATTTCACTTAACATAATGACTTCCTGTTCCATCCATGTTGCTGCAAATGACAGGA
TTTCGTTCTTAATTTCAATTAAAATAACCACACATGGCAAAAA

386/615

FIGURE 382

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQ
PVKGHGTLGESPMPPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 383

TTCTGAAGTAACGGAAGCTACCTTGTATATAAAGACCTCAACACTGCTGACC**ATG**ATCAGCGCAG
CCTGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTTCAAGTAGCACCTCTATCAG
TTATGGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTTCATTTACTGTAATGATC
GCTTTCTGACATCCATTCCAACAGGAATACCAAGGATGCTACAACCTCTCTACCTTCAGAAAC
ACCAAATAAATAATGCTGGGATTCTTTCAGATTTGAAAAAAGTTGCTGAAAGTAGAAAGAATAT
ACCTATACCACAACAGTTTTAGATGAATTTCTTACCAACCTCCCAAAGTATGTAAAAGAGTTAC
ATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCCTATCTGG
AAGAATTACATTTAGATGACAACTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCATTCCGAG
ACAGCAACTATCTCCGACTGCTTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTT
TGCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCAT
CTCTTCAAGGTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATG
GTTTAGGTGACAAAGTTTTCTTCAACCTAGTTAATTTGACAGAGCTGTCCTGTTGCGGAATT
CCCTGACTGCTGCACCAGTAAACCTTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATA
ACCACATCAATCGGGTGCCCCCAAATGCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATA
TGTCCTAATAATAACCTAAGTAATTTACCTCAGGGTATCTTTGATGATTTGGACAATATAACAC
AACTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT
TACAATCACTACCTGTGAAGGTCAACGTGCGTGCGGCTCATGTGCCAAGCCCCAGAAAAGGTT
GTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAA
GCACCATTCAGATAAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAG
CTCCAGTGACCAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG
GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCATA
TCTCTTGGAACCTTGCTCTACCTATGACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATA
GCCCCGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAACGCAGTGAGTACTTGGTCA
CAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTCCCATGGAAACCAGCAACCTCT
ACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTTCGAATGTACAACC
CTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACCTTTGG
CTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGT
ATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAACTGTGCATATAGCAAAGGGAGGAGAAGAA
AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTT
CTTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCA
TATTTCTCTAATGGAATGAATCTGTACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAA
GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCA**TGA**TGCTGAAGGACTCACA
GCAGACTTGTGTTTTTGGGTTTTTTTAAACCTAAGGGAGGTGATGGT

FIGURE 384

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRLFSTIPTGIPEDATTLYLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIIRTITYDSLKIPYLEELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPVKVNVVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPDIKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITETIVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVCIIETETAPLRMYNPPTTTLNREQEKEPYKNPNLPLAAIIGGAVALVTIALALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNSILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSES SSNRSYRDSGIPDSDHSHS

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 385

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGG
 CAGCCCTGCCGGGGCCACTTGTCTTC**ATGT**CTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAG
 GGCGTGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGA
 GCAGAGGCCCGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACC
 GCTGTGGTCCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT
 GATCCTGTCTGCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCT
 GGATGGCATGCTGGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGGAGAAGTGGGC
 CCAGGAGCCCCCTGCTGCAGCCGCTGAGCCTGCGCGTGCGGATGCTGGGGGAGAAGCTGGAGGC
 TGCCATCCAGAGATCCCTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCA
 GCTGACCCTCCAGCCCCGGGTTTTGGAAGCTCCACATGCCTGGATCCCACTGATGCCTCCTT
 GGTGTACCCACGTTTCGGGCCCCAGGACTCATTTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT
 GCAGCTGCTGGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAG
 CCTCATGACCAAGCCCCGGCTGCTCAGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCTCTG
 GGCCAGAATGAGGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCAACCTCTT
 CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATCGGATACGCCTACCCTACCCG
 GGACATCTTCATGGAACATCATGTTCTGTGGAATGGGCGGCTTCTCCGACTTCTACAAGCT
 CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAAACAGCAGGAAGGATGCTTCGGGGAGCCTGA
 TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTTTTCGAGGAGAGTGAA
 GAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGTTGCTCAGGCTGGAGTACAGTGGCGCAA
 TCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGCAATTCTCTTGCCTCATCCTCCCGAG
 TAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAATTTTTTATATTTTTTTAGTAGAGAC
 AGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGATCTCAAGAGATCCGCCCCACCTC
 AGGCTCCCAAAGTGTGGGATTA**TAG**GTGTGAGCCACCGTGTCTGGCTGAAAAGCACTTTCAAA
 GAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTCATGGGGGCTCTCTCCCC
 TAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGGGTGGCTTCTCTATACA
 TCCTGGCAGAATACCCCCCAGCAACAGAGAGCCACACCCATCCACACCGCCACCACCAAGCA
 GCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTTAGTCCTCA
 TCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGGATAAGCAAA
 GCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGGGGCCGGG
 AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAAAA

390/615

FIGURE 386

MSARGRWEGGRRACRGS LG LARAQGAERVTSSEQRPAMASLG LLLLLLLLTALPPLWSSSLPG
LDTAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPL
SLRVGMLGEKLEAAIQ RSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQ
DSFSEERSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQ
GPLQQSQDYINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILS
WQKQQEGCFGE PD AEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLP
PGFKQFSC LILPSSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

20250101 09:55:01

391/615

FIGURE 387

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGC
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCCGGGGCTGCTGCTGAGGGATCGG
GAGGGAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAGTTGTTCCAGGGG
TGAAGCCTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT
TCCTTAAGACAGATGGGAGTTTGTGGTTTCATGATATACCTTCTGGATCTTATGTAGTGGAAG
TTGTATCTCCAGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGA
GAGCAAGATATGTGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAA
TGAAATCTTCAGGTCCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTC
TAATGAACCCAATGGTTATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAG
TGGTCAACACAAGTGATCCTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATT
CCAACCATGAGTTGCCTGATGTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTG
GCAAATCTAGCAGCGGCAGCAGTAAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAG
GCCGTCCAGAGCTGGCATTGTCACAAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAA
CCGTGTGAAGCAACTACTATAAACTTGAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGT
ATGAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTTATAGAAAATGCCATTAATAAA
TTATATGAACACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAA
AAAAAAAAAAAAAAAAAAAAA

10053536.041502

FIGURE 388

MAAALWGFFPVLALLLLSGDVSSEVPGAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWIS
AARVLVDGEEHVGFLKTDGSFVVDI PSGSYVVEVVS PAYRFD PVRVDITSGKMRARYVNYI
KTSEVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLIFVLLPKVNTSDPD
MRREMEQSMNMLNSNHELDPVSEFMTRLFSSKSSSGKSSSGSKTGKSGAGKR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 389

GTCGTGTGCTTGGAGGAAGCCGCGGAACCCCCAGCGTCCGTCCTCC**ATG**GCGTGGAGCCTTGGGAG
 CTGGCTGGGTGGCTGCCTGCTGGTGTGTCAGCATTGGGAATGGTACCACCTCCCGAAAAATGTCAG
 AATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTGCTTTTGCCAAAGGGAA
 CCTGACTTTTCACAGCTCAGTACCTAAGTTATAGGATATTCCAAGATAAATGCATGAATACTAC
 CTTGACGGAATGTGATTTCTCAAGTCTTTCCAAGTATGGTGACCACACCTTGAGAGTCAGGGC
 TGAATTTGCAGATGAGCATTGAGACTGGGTAAACATCACCTTCTGTCCTGTGGATGACACCAT
 TATTGGACCCCCCTGGAATGCAAGTAGAAGTACTTGCTGATTCTTTACATATGCGTTTCTTAGC
 CCCTAAAATTGAGAATGAATACGAACTTGAGCTATGAAGAATGTGTATAACTCATGGACTTA
 TAATGTGCAATACTGGAAAAACGGTACTGATGAAAAGTTTCAAATTACTCCCCAGTATGACTT
 TGAGGTCCTCAGAAACCTGGAGCCATGGACAACCTTATTGTGTTCAAGTTCGAGGGTTTCTTCC
 TGATCGGAACAAAGCTGGGGAATGGAGTGAGCCTGTCTGTGAGCAAACAACCCATGACGAAAC
 GGTCCCCCTCCTGGATGGTGGCCGTATCCTCATGGCCTCGGTCTTCATGGTCTGCCTGGCACT
 CCTCGGCTGCTTCTCCTTGCTGTGGTTCGTTTACAAGAAGACAAAGTACGCCTTCTCCCCTAG
 GAATTCTCTTCCACAGCACCTGAAAGAGTTTTTGGGCCATCCTCATCATAACACACTTCTGTT
 TTTCTCCTTTCCATTGTGCGATGAGAATGATGTTTTTGACAAGCTAAGTGTGATTGCAGAAGA
 CTCTGAGAGCGGCAAGCAGAATCCTGGTGACAGCTGCAGCCTCGGGACCCCGCCTGGGCAGGG
 GCCCCAAAGCT**TAG**GCTCTGAGAAGGAAACACACTCGGCTGGGCACAGTGACGTACTCCATCTC
 ACATCTGCCTCAGTGAGGGATCAGGGCAGCAAACAAGGGCCAAGACCATCTGAGCCAGCCCCA
 CATCTAGAACTCCAGACCTGGACTTAGCCACCAGAGAGCTACATTTTAAAGGCTGTCTTGGCA
 AAAATACTCCATTTGGGAACTCACTGCCTTATAAAGGCTTTCATGATGTTTTTCAGAAGTTGGC
 CACTGAGAGTGTAATTTTTCAGCCTTTTATATCACTAAAATAAGATCATGTTTTTAATTGTGAGA
 AACAGGGCCGAGCACAGTGGCTCACGCCTGTAATACCAGCACCTTAGAGGTCGAGGCAGGCGG
 ATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAATATGGTGAAACCCAGTCTCTACTAA
 AAATACAAAAATTAGCTAGGCATGATGGCGCATGCCTATAATCCCAGCTACTCGAGTGCCTGA
 GGCAGGAGAATTGCATGAACCCGGGAGGAGGAGGAGGAGGTTGCAGTGAGCCGAGATAGCGGC
 ACTGCACTCCAGCCTGGGTGACAAAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAATTGTG
 AGAAACAGAAATACTTAAAATGAGGAATAAGAATGGAGATGTTACATCTGGTAGATGTAACAT
 TCTACCAGATTATGGATGGACTGATCTGAAAATCGACCTCAACTCAAGGGTGGTCAGCTCAAT
 GCTACACAGAGCACGGACTTTTGGATTCTTTGCAGTACTTTGAATTTATTTTTCTACCTATAT
 ATGTTTTATATGCTGCTGGTGCTCCATTAAAGTTTTACTCTGTGTTGC

394/615

FIGURE 390

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83551
><subunit 1 of 1, 325 aa, 1 stop
><MW: 37011, pI: 5.09, NX(S/T): 4
MAWSLGSWLGGCLLVSALGMVPPPENVRMNSVNFKNILQWESPAFAKGNLTFTAQYLSYRIFQ
DKCMNTTTLTECDFSSLSKYGDHTLRVRAEFADHSDWVNITFCPVDDTIIGPPGMQVEVLADS
LHMRFLAPKIENEYETWTMKNVYNSWTYNVQYWKNGTDEKFQITPQYDFEVLRLNLEPWTTYCV
QVRGFLPDRNKAGEWSEPVCEQTTHDETVPSPWMVAVILMASVFMVCLALLGCFSLWCVYKKT
KYAFSPRNSLPQHLKEFLGHPHNTLLFFSFPLSDENDVFDKLSVIAEDSESGKQNPBGDSCSL
GTPPGQGPQS
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 222-245

N-glycosylation sites.

amino acids 49-53, 68-72, 102-106, 161-165

N-myristoylation sites.

amino acids 6-12, 316-322

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAAGGTTCTGACGCGATGAGGA
AGCACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGG
TCCAGACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCCAGCACTGCC
AGATCACTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGC
TCGACATTGACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCAACGGCT
GCATCAATGCCACCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACC
AGCAGGTGCTCTGGCGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTTGGTTGG
AGAGGGGGCGCAGGACTTCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGA
TCTGGCTCATGGTGAAATTAAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAA
TCCTGGCAAGTGACCCAGCTCTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGC
GGCGATGCACTCGCACTGCAAATGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCT
GATAGATGGGGGACTGTGGCTTCTCCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCA
CACTAGATTAGTAGTAAATGCTTGATGAGAAGAACACATCAGGCACTGCGCCACCTGCTTCAC
AGTACTTCCCAACAACCTCTTAGAGGTAGGTGTATTCCCGTTTTTACAGATAAGGAACTGAGGC
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGAGCCAGGATTCAAC
CCTGGCTTGTCTAACCCCAGGTTTTTCTGCTCTGTCCAATTCCAGAGCTGTCTGGTGATCACTT
TATGTCTCACAGGGACCCACATCCAAACATGTATCTCTAATGAAATTGTGAAAGCTCCATGTT
TAGAAATAAATGAAAAACCTGA

396/615

FIGURE 392

MRKHLSSWWLATVCMLLFSLHSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQG
RKLDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNK
LHQQVLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

Downloaded from www.sciencedirect.com

FIGURE 393

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGAA
TGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCCCT
CAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGCCAA
TTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTGAGAG
ATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTCCTTTG
CTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTTACCCTGCTTCTC
TGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCACCTTTGGGCCCTGCCAAGTGAAGGGG
GTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAGCTCAGGAT
AACATCACGAGTCCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGATGCTGAGAGC
TGTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTGTTTTCAAAAACCAACAAT
AGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAACCTTTGTTCTCATC
GTGTCACAACCTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGACAGTGCACACAGG
CGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGCTCTGACCAAAGCC
CTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGCTCT**TGA**ATGTCTAGA
CCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCAATTTCAAACAGTCTCCCTTCC
TATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTCTTGGCCCAGGATTATT
GTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAGGTGCCTCTGGATGCTGT
GAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTTAATTAATGTCAGTATTTT
AACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGGCAGCAGAATATTGTGCCCC
ATGCTTCTTTACCCCTCACAAATCCTTGCCACAGTGTGGGGCAGTGGATGGGTGCTTAGTAAGT
ACTTAATAAACTGTGGTGCTTTTTTTTGGCCTGTCTTTGGATTGTTAAAAAACAGAGAGGGATG
CTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACTGTCTTCTGATATCTGCAGGGAC
AGAGCATTGGGGTGGGGGTAAAGGTGCATCTGTTTGAAAAGTAAACGATAAAATGTGGATTAAA
GTGCCCAGCACAAAGCAGATCCTCAATAAACATTTCAATTTCCACCCACACTCGCCAGCTCAC
CCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTTATCCTAGTCATTCTTCCCTAATCT
TCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGACATTGCACCTGGATGTACTATCCAATC
TGTGATGACATTCCCTGCTAATAAAAGACAACATAACTCCAAAAAAAAAAAAAAAAAAAAA
AAAA

398/615

FIGURE 394

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002
><subunit 1 of 1, 206 aa, 1 stop
><MW: 23799, pI: 9.12, NX(S/T): 3
MNFQQRLQSLWTLARPFPCPLLATASQMQMVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVKG
VVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLLEFYLKTVFKNHHN
RTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAALTKA
LGEVDILLTWMQKFYKL
```

Signal sequence:
amino acids 1-42

N-glycosylation sites.
amino acids 85-89, 99-103, 126-130

1005555 01503
20510 98525001

399/615

FIGURE 395

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCAT
TTCCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTC
CTCTCTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCA
AGCTGGAAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTT
TGAAATTTCAACTTTCAGATTTCAGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCAT
GATGCTGAGGTTTGGGGT

10053536 0.1503
20510 9825001

400/615

FIGURE 396

MFRSSLLFWPPLCLLSLFLILISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT
CEGLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

THE UNIVERSITY OF CHICAGO

FIGURE 397

CATGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTG
TGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTT
CATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGAACCTACTG
TGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTG
GGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCCTATCCTTTTTTTAGAAACACAAAT
TGGCCCAACAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCT
GAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCACTGTGTGACCAA
CCACACGCTGGTGTCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTT
CGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGA
TCAATCATCAGAGTTCAAGGCTAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGT
GTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACA
CCCAGCAAATTTGATTTTATGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGA
AAAAATCGTGATTAACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGA
TATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAA
CCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGA
AATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAG
CAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCACTGACAT
TTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATT
ATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCC
TCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGA
AGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATTCTTCGCT
GTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGA
GGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAAC
CTATCTCATGCAATTCATGGAGGAATGGGGGTATATATGTGCAGATGGAAAAC**TGAT**GGCCAACA
CTTCCTTTTTGCCTTTTTGTTTCCTGTGCAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAA
GTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTCAAGTGTCTGTGAGAA

FIGURE 398

MPLPPLLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQYF
IYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFLETQI
GPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTWSQCVTN
HTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQSSEFKAKIIFWYVLPISITV
FLFSVMGYSIYRYIHVGKEKHPANLILYGNFEDKRFFVPAEKIVINFITLNISSDDSKISHQD
MSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENTEGTSLTQQESLS
RTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLSQALAVLGPQTLQYSYTP
QLQDLDPPLAQEHTDSEEGPEEPPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEE
GLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 240-260

N-glycosylation sites.amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192,
304-307, 523-526**Tyrosine kinase phosphorylation site.**

amino acids 385-392, 518-526

N-myristoylation sites.

amino acids 53-58, 106-111, 368-373, 492-497

Tissue factor

amino acids 1-278

FIGURE 399

CCGGCG**ATG**TCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAG
CCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA
ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACCTAGTGTGCAACAGGGGAC
TATTCAATTTTGTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAG
GCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAAT
TACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATC
GGCTTCCCTGTAGAGCTGAACACAGTCTATTTTCATTGGGGCCCATAATATTCCTAATGCAAAT
ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATA
ATGAAATATAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGT
AAGAAGAATGAGGAGACAGTAGAAGTGAAGTTTACAACCACTCCCTGGGAAACAGATACATG
GCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAA
CAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTG
ACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGC
CCACAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCT
CTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGG
AGGCACGAAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCATTAAGGTT
CTTGTGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTT
CAAAACCATTGCAGAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATG
GGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCC
AATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCAGTGAGAAC
TCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATTTCAT
CTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAATGCTCTC
AGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCAT
GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAAGCCTGCCACGATGGCTGCTGCTCCTTG
TAG

FIGURE 400

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRLDLRVEPVTTTSVATGDYS
ILMNVSWSVLRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGF
PVELNTVYFIGAHNIPNANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKK
NEETVEVNFTTTTPLGNRYMALIQHSTIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTP
YFPTCGSDCIRHKGTVVLCPOGTGVPFPLDNNKSKPGGWLPLLLLSLLVATWVLVAGIYLMWRH
ERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHCRSEVILEKWQKKKIAEMGP
VQWLATQKKAADKVVFLLSNDVNSVCDGTGCGKSEGSPSENSQDLFPLAFNLFCSDLRSQIHLH
KYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQVVSAGKRSQACHDGCCSL

Important features of the protein:**Signal peptide:**

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201
and 283 - 287**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

GGGAACACAGGAACTATCAGCCCGTCGGCCTCCGGGCCCTGCATTCTCTAGCCA**ATGGACCG**
GGACCTTTTTCGGCAGTCGCTAAATTGCCACGGGTCGTCTTTGCTCTCTCTACTTTCGGAG
CGAACAGCAGGACAATCCACACTTCCGTAGCCTCCTGGGGTTCGGCCGCCGAGCCAGCCCCG
GGGCCCGCCGCCCCAGCACCCGTTTCAGGGGCAGAAAAGAGAAGAGAGTTGACAACATCGA
GATACAGAAATTCATCTCCAAAAAAGCGGATCTGCTTTTTTGCACTTTCTTGGAAATCAGA
TGCACCTGCAACTTCTGAAATTAATGAAGACAGTGAAGATCATTATGCAATCATGCCACC
TTTAGAGCAATTCATGGAGATACCTAGTATGGATCGGAGAGAGCTGTTTTTCCGAGATA
TGAGCGTGGTGATATAGTGATTGGAAGAATTAGTTCTATTTCGGGAATTCGGTTTTTTTCAT
GGTGTTGATCTGTTTAGGAAGTGGTATCATGAGAGATATAGCCCACTTAGAAATCACAGC
TCTTTGTCCCTTAAGAGATGTGCCTTCTCACAGTAACCATGGGGATCCTTTATCATATTA
CCAAACTGGTGACATCATTCGAGCTGGAATCAAGGATATTGACAGATACCATGAAAAGCT
AGCAGTATCTCTGTATAGCTCTTCTCTTCCACCACACCTATCTGGTATTAAATTAGGTGT
AATTAGCTCTGAAGAGCTTCCCTTTATACTACAGGAGAAGTGTTGAGCTAAATAGCAATTC
TTTGAGTCCCTATGAAAATGTCATGCAGAGTTCCTTGGGATTTGTTAATCCAGGAGTAGT
TGAATTCCTTCTAGAAAAACTAGGAATAGATGAATCTAATCCACCATCTTTAATGAGAGG
CCTACAAAGCAAAAATTTCTCTGAAGATGATTTTGGCTTCTGCATTGAGAAAAAACAATC
CGCATCTTGGGCTTTAAAATGTGTGAAGATCGGAGTTGACTATTTTAAAGTTGGACGCCA
TGTGGATGCTATGAATGAATACAATAAAGCTTTGGAAATAGACAAACAAAACGTGGAAGC
TTTGGTAGCTCGTGGAGCATTATATGCGACAAAAGGAAGTTTGAACAAAGCAATAGAAGA
TTTTGAGCTTGCATTAGAAAACGTGCCAACTCACAGAAATGCAAGAAAATACCTCTGCCA
GACACTTGTAGAGAGAGGAGGACAGTTAGAAGAAGAAGAAAAGTTTTTAAATGCTGAAAG
TACTATAAGAAAGCCTTGGCTTTGGATGAGACTTTTAAAGATGCAGAGGATGCTTTGCA
GAAACTTTCATAAATATATGCAGAAATCTTTGGAATTAAGAGAAAAACAAGCTGAAAAGGA
AGAAAAGCAGAAAACAAAGAAAATAGAAACAAGTGCAGAAAAGTTGCGTAAGCTCTTAAA
AGAAGAGAAGAGGCTAAAGAAGAAAAAGAAAGAAATCAACTTCTTCTTCAAGTGTTTTCTC
TGCTGATGAATCAGTGTCTTCATCATCATCCTCTTCTCTCTTCTGGTCACAAAAGGCATAA
GAAACATAAGAGGAACCGTTCAGAGTCTTCTCGCAGTTCCAGAAGGCATTTCATCTAGGGC
ATCCTCAAATCAGATAGATCAGAATAGGAAAAGATGAGTGCTACCCAGTTCCAGCTAATAC
TTCAGCATCTTTTTCTTAACCATAAAACAAGAAAGTGGAGAACTACTGGGGAAGCAGGATAG
GTTACAGTATGAAAAGACACAGATAAAAGAGAAAGATAGATGCCCTCTCTCTTCATCTTC
ACTTGAAATACCGGATGATTTTGGAGTGTACTCCTATTTATTTAAAAAGTTAACTATAAA
ACAGCCTCAGGCAGGTCCTTCAGGAGATATTCCAGAAGAGGGCATTGTTATCATAGATGA
CAGCTCCATTTCATGTTACTGACCCTGAAGACCTTCAAGTGGGACAAGATATGGAGGTGGA
AGACAGTGGTATTGATGATCCTGACCACGGG**TAGG**CTTAGGTTTATGTGTGTGTATGTGT
CTTAGTTTTTAAACAAAAAAATTAAGTAAAAAACTAAAAATAGAAAAATGCTTAGAG
AATAAGGATATAAAGAATATTTTTGTGCAGTTGAACAATGAGTGCTTAAGCTAAATGTCA
TCACAAAAGAGTAAAAAAATTTTACAAAATTAAGTAAAGTTAAAAAGCTCTAGG
AAGCTAAGGTCAATTTATTATTGGAGAAATAAAATTAATTTTTATGAATTTACTGT

406/615

FIGURE 402

MDRDLLRQSLNCHGSSLLSLLRSEQQDNPHFRSLLGSAAEPARGPPPQHPLQGRKEKRVD
NIEIQKFISKKADLLFALSWKSDAPATSEINEDSEDHYAIMPPLEQFMEIPSMRRELFF
RDIERGDIVIGRISSIREFGFFMVLIICLGSGIMRDIAHLEITALCPLRDVPSHSNHGDPL
SYYQTGDIIRAGIKDIDRYHEKLAVSLYSSSLPPLHSGIKLGVISSEELPLYRRSVELN
SNSLESYENVMQSSSLGFVNPGVVEFLLEKLGIDESNPPSLMRGLQSKNFSEDDFASALRK
KQSASWALKCVKIGVDYFKVGRHVDAMNEYNKALEIDKQNVREALVARGALYATKGSINKA
IEDFELALENCPTHRNARKYLCQTLVERGGQLEEEEEKFLNAESYYKKALALDETFKDAED
ALQKLHKYMQKSLELEKQAEKEEKQKTKKIETSAEKLRLKLLKEEKRLKKRRKSTSSSS
VSSADESVSSSSSSSSSGHKKRHHKRNRSSESSRRHSSSRASSNQIDQNRKDECYPVP
ANTSASFLNHKQEVKLLGKQDRLQYEKTIKEKDRCPLSSSSLEIPDDFGVYSYLFKKL
TIKQPQAGPSGDIPEEGIVIIDDSSIHVTDPEDLQVGQDMEVEDSGIDDPDHG

Important features of the protein:

Signal peptide:

Amino acids 1-23

Transmembrane domain:

Amino acids 138-155

N-glycosylation sites:

Amino acids 288-292;508-512;542-546

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 300-304;472-476;473-477;517-521;598-602

N-myristoylation sites:

Amino acids 218-224;222-228;271-277;348-354

Amidation site:

Amino acids 52-56

Cell attachment sequence:

Amino acids 125-128

406/615

407/615

FIGURE 403

CCGAGGCGGGAGGAGCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTTC
CAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATA
GGAAAATAACTTGGGATTTTATATTGGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTT
ATGACAAACTTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAG
AGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCC
CGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCA
CTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCTGGAGCTC
ACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAAGAAGTACA
TGTCAGAAAATAAGGGAGTTCTCTGCATGGGGGTGATGAAGACAGACCTTTCCAGACTTTG
ACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCATTCTGCCAACTGCACTG
GCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAAATTTGAGAGGC
TCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTCAGCATTTTTTGT
GCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAGTGGTGGCGCTGCT
TTCCTGAGCGGTGGTTCCCATTTTCTTATCCATGGAGGAGACCTCTGAACAGATCACAAATGT
TACGTGAGCTTTTTCTCTGTTTTCACTCACCTGCCATTTCCAAAAGATGCCTCTTTAAACAAGT
GCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAGATGCCTGACCTATTTA
TCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCAGTGCCGAAGACATTGTC
AGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTGACACCACCCACTGGAAGG
TCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGATGGAACCGCTTTCTCAGAAC
TGTAGGAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGAAAACCAGGTTGAAAGGGGAA
AAATAAAAACAAAACGATGAACTGCAAAAA

408/615

FIGURE 404

MDLAANEISYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVYK
VLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPLHG
GDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKTGKP
LLEEEIQHFQYPEATEGFSEGGFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFVFTHL
PFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQRRHCQSVAMPIEPGD
IGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

408/615

409/615

FIGURE 405

TGCCGGGCTGCGGGGCGCCTTGACTCTCCCTCCACCCTGCCTCCTCGGGCTCCACTCGTCTGCCCTGGACTCCC
GTCTCCTCCTGTCTCCGGCTTCCCAGAGCTCCCTCCTTATGGCAGCAGCTTCCCGCGTCTCCGGCGCAGCTTCT
CAGCGGACGACCCTCTCGCTCCGGGGCTGAGCCCAGTCCCTGGATGTTGCTGAAACTCTCGAGATCATGCGCGGG
TTTGGCTGCTGCTTCCCCGCCGGGTGCCACTGCCACCGCCGCCCTCTGCTGCCGCCGTCCGCGGGATGCTCAG
TAGCCCCGTGCCCCGGCCCCCGCGATCCTGTGTTCCCTCGGAAGCCGTTTGCTGCTGCAGAGTTGCACGAAC TAGTC
ATGGTGCTGTGGGAGTCCCCGCGGCAGTGCAGCAGCTGGACACTTTGCGAGGGCTTTTGCTGGCTGCTGCTGCTG
CCCGTCATGCTACTCATCGTAGCCCGCCCGGTGAAGCTCGCTGCTTTCCCTACCTCCTTAAGTGACTGCCAAACG
CCCACCGGCTGGAATTGCTCTGGTTATGATGACAGAGAAAATGATCTCTTCCCTCTGTGACACCAACACCTGTAAA
TTTGATGGGGAATGTTTAAGAATTGGAGACACTGTGACTTGCGTCTGTGCTCAGTTCAAGTGCAACAATGACTATGTG
CCTGTGTGTGGCTCCAATGGGGAGAGCTACCAGAATGAGTGTACCTGCGACAGGCTGCATGCAAACAGCAGAGT
GAGATACTTGTGGTGTGAGAAGGATCATGTGCCACAGATGCAGGATCAGGATCTGGAGATGGAGTCCATGAAGGC
TCTGGAGAACTAGTCAAAAGGAGACATCCACCTGTGATATTTGCCAGTTTGGTGCAGAATGTGACGAAGATGCC
GAGGATGTCTGGTGTGTGTGTAATATTGACTGTTCTCAAACCAACTTCAATCCCCCTCTGCGCTTCTGATGGGAAA
TCTTATGATAATGCATGCCAAATCAAAGAAGCATCGTGTGAGAAACAGGAGAAAATTGAAGTCATGTCTTTGGGT
CGATGTCAAGATAACACAAC TACAAC TACTAAGTCTGAAGATGGGCATTATGCAAGAACAGATTATGACAGAGAAT
GCTAACA AATTAGAAGAAAGTGCCAGAGAACACCACATACCTTGTCCGGAACATTACAATGGCTTCTGCATGCAT
GGGAAGTGTGAGCATTCTATCAATATGCAGGAGCCATCTTGCAAGGTGTGATGCTGGTTATACTGGACAACACTGT
GAAAAAAGGACTACAGTGTTCTATACGTTGTTCCCGGTCCCTGTACGATTTTCAGTATGTCTTAATCGCAGCTGTG
ATTGGAACAATT CAGATTGCTGT CATCTGTGTGGTGGTCCCTCTGCATCACAAGGAAATGCCCCAGAAGCAACAGA
ATTCACAGACAGAAGCAAAATACAGGGCACTACAGTTCAGACAATACAACAAGAGCGTCCACGAGGTTAATCTAA
AGGGAGCATGTTTCACAGTGGCTGGACTACCGAGAGCTTGACTACACAATACAGTATTATAGACAAAAGAATAA
GACAAGAGATCTACACATGTTGCCTTGCATTTGTGGTAATCTACACCAATGAAAACATGTACTACAGCTATATTT
GATTATGTATGGATATATTTGAAATAGTATACATTGTCTTGATGTTTTTCTGTAATGTAAATAAACTATTTATA
TCACACAATATAGTTTTTTCTTTCCCATGTATTTGTATATATAATAAATACTCAGTGATGAG

20510-9852501

[illegible]

MVLWESPRQCSSWTECEGFCWLLLLLPMLLIVARPVKLAAPFTSLSDCQTPGTGWNCSGY
DDRENDLFLCDTNTCKFDGECLRIGDVTVCVCQFKCNNDYVPVCGSNGESYQNECYLRQ
AACKQQSEILVSEGSCATDAGSGSGDGVHEGSGETSQKETSTCDICQFGAECDDEDAED
VWCVCNIDCSQTNFNPLCASDGKSYDNACQIKEASCQKQEKIEVMSLGRCDNTTTTTTK
SEDGHYARTDYAENANKLLEESAREHHIPCPEHYNGFCMHGKCEHSINMQEPSCRCDAGY
TGQHCYCKTDYSVLVYVPGVPFRQYVYLLAAVIGTIQIAVICVVVLCITRKCPRSNRIHRQ
KQNTGHYSSDNTTTRASTRLI

FIGURE 407

CTCGCAGCCGAGCGCGGCCGGGAAGGGCTCTCCTTCCAGCGCCGAGCACTGGGCCCTGGCAG
 ACGCCCCAAGATTGTTGTGAGGAGTCTAGCCAGTTGGTGAGCGCTGTAATCTGAACCAGCTGT
 GTCCAGACTGAGGCCCCATTTGCATTGTTTAACTACTTAGAAAATGAAGTGTTTCATTTTTAA
 CATTCTCTCTCCAATTGGTTTAAATGCTGAATTACTGAAGAGGGCTAAGCAAAACCAGGTGCTT
 GCGCTGAGGGCTCTGCAGTGGCTGGGAGGACCCCGCGCTCTCCCCGTGTCTCTCCACGACT
 CGCTCGGCCCCCTCTGGAATAAAACACCCGCGAGCCCCGAGGGCCCAGAGGAGGCCGACGTGCC
 CGAGCTCCTCCGGGGTCCCGCCCCGCGAGCTTTCTTCTCGCCTTCGCATCTCCTCCTCGCGCG
 TCTTGACATGCCAGGAATAAAAGGATACTCACTGTTACCATTTCTGGCTCTCTGTCTTCCAA
 GCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAGGACAGT
 GTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGACATGATGTGTGTTA
 ACCAAAATGGCGGGTATTTATGCATTCCCCGGACAAACCCTGTGTATCGAGGGCCCTACTCGA
 ACCCTACTCGACCCCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAA
 ACTATCCACGATCTCCAGGCCTCTTATATGCCGCTTTGGATACCAGATGGATGAAAGCAACC
 AATGTGTGGATGTGGACGAGTGTGCAACAGATTCCCACCAGTGCAACCCCCACCCAGATCTGCA
 TCAATACTGAAGGCGGGTACACCTGCTCCTGCACCGACGGATATTGGCTTCTGGAAGGCCAGT
 GCTTAGACATTGATGAATGTGCTATGGTTACTGCCAGCAGCTCTGTGCGAATGTTCTCTGGAT
 CCTATTCTTGTACATGCAACCCTGGTTTTTACCCTCAATGAGGATGGAAGGTCTTGCCAAGATG
 TGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAAACCTGCGTCAACACCTACGGCTCTCTCA
 TCTGCCGCTGTGACCCAGGATATGAACTTGAGGAAGATGGCGTTTCATTGCAGTGATATGGACG
 AGTGCAGCTTCTCTGAGTTCTCTGCCAACATGAGTGTGTGAACCAGCCCCGGCACATACTTCT
 GCTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAAGCTGCCAAGACATCAACGAAT
 GTGAGCACAGGAACACACGTGCAACCTGCAGCAGACGTGCTACAATTTACAAGGGGGCTTCA
 AATGCATCGACCCCCATCCGCTGTGAGGAGCCTTATCTGAGGATCAGTGATAACCGCTGTATGT
 GTCTGTGAGAACCCCTGGCTGCAGAGACCAGCCCTTTACCATCTTGTACCGGGACATGGACG
 TGGTGTGAGGACGCTCCGTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCCTG
 GGGCCTATTACATTTTCCAGATCAAATCTGGGAATGAGGGCAGAGAATTTTACATGCGGCAAA
 CGGGCCCCATCAGTGCCACCCTGGTGATGACACGCCCCATCAAAGGGCCCCGGGAAATCCAGC
 TGGACTTGGAATGATCACTGTCAACACTGTCATCAACTTCAGAGGCAGCTCCGTGATCCGAC
 TGCGGATATATGTGTCGAGTACCCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTGCCTCT
 CATTGGCACCAAGGGACAGGAGAAGAGAGGAAATAACAGAGAGAATGAGAGCGACACAGACGT
 TAGGCATTTCTGTGAACGTTTCCCCGAAGAGTCAGCCCCGACTTCTGACTCTCACCTGTA
 CTATTGCAGACCTGTCACCCTGCAGGACTTGCCACCCCCAGTTCCTATGACACAGTTATCAAA
 AAGTATTATCATTGCTCCCCTGATAGAAGATTGTTGGTGAATTTTCAAGGCCTTCAGTTTATT
 TCCACTATTTTCAAAGAAAATAGATTAGGTTTGCGGGGGTCTGAGTCTATGTTCAAAGACTGT
 GAACAGCTTGCTGTCACTTCTTCACTCTTCCACTCCTTCTCTCACTGTGTTACTGCTTTGCA
 AAGACCCGGGAGCTGGCGGGGAACCTGGGAGTAGCTAGTTTGCTTTTTGCGTACACAGAGAA
 GGCTATGTAAACAAACCACAGCAGGATCGAAGGGTTTTTAGAGAATGTGTTTCAAACCATGC
 CTGGTATTTTCAACCATAAAAGAAGTTTTCAGTTGTCCTTAAATTTGTATAACGGTTTAATTCT
 GTCTTGTTTCATTTTGAAGTATTTTAAAAAATATGTCGTAGAATTCCTTCGAAAGGCCTTCAGA
 CACATGCTATGTTCTGTCTTCCCAAACCCAGTCTCCTCTCCATTTTAGCCCAGTGTTTTCTTT
 GAGGACCCCTTAATCTTGCTTTCTTTAGAATTTTACCCAATTGGATTGGAATGCAGAGGTCT
 CCAAACCTGATTAAATATTTGAAGAGA

412/615

FIGURE 408

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQN
GGYLCIPRTNPVYRGPYSNPYSTPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDESNQCV
DVDECATDSHQCNPTQICINTEGGYTCSCTDGYWLLEGQCLDIDECRYGYCQQLCANVPGSYS
CTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGS LICRCDPGYELEEDGVHCSDMDECS
FSEFLCQHECVNQPGTYFCSCPPGYILLDDNRSCQDINECEHRNHTCNLQQTCYNLQGGFKCI
DPIRCEEPYLRISDNRCMCPAENPGCRDQPFTILYRDMDVVSGRSVPADIFQMQATTRYPGAY
YIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPREIQLDLEMITVNTVINFRGSSVIRLRI
YVSQYPF

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 283-287, 296-300

N-myristoylation sites.

amino acids 21-27, 64-70, 149-155, 186-192, 226-232, 242-248,
267-273, 310-316

Aspartic acid and asparagine hydroxylation sites.

amino acids 144-156, 181-193, 262-274

Cell attachment sequence.

amino acids 54-57

Calcium-binding EGF-like.

amino acids 131-166, 172-205, 211-245, 251-286

FIGURE 409

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGCTCA
GGGAGGAGCACCGACTGCGCCGCACCCTGAGAGATGCTTGGTGCCATGTGGAAGGTGATTGTT
TCGCTGGTCCCTGTTGATGCCTGGCCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAAGTGT
TCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGCTGGG
AAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCAGGACTGAACATGAAGAGT
TATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTCCTCCCA
GCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGGAGGTTCA
TCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCAACAAGTAACATGACCTTGCCT
GACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCAGTGGGCACA
GGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGTAGCACGGGAT
TTATACAGTGCCTAATTGAGTTTTCAGATATTTCTGAATATAAAAAATAATGACTTTTAT
GTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTCATCCATTCCCTC
AACCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGATGGATATTCTGAT
CCCGAATCAATTATAGGGGGCTATGCAGAATTCTGTACCAAATTGGCTTGTGATGAGAAG
CAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATCAGGAAGCAGAACTGG
TTTGAGGCCTTTGAAAATACTGGATAAACTACTAGATGGCGACTTAACAAGTGATCCTTCTTAC
TTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGTGCACGGAACCTGAGGAT
CAGCTTTACTATGTGAAATTTTTGTCACTCCAGAGGTGAGACAAGCCATCCACGTGGGGAAAT
CAGACTTTTAATGATGGAATATAGTTGAAAAGTACTTGCGAGAAGATACAGTACAGTCAGTT
AAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGATCTACAATGGCCAACTGGAC
ATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGCATGGACTGGAAAGGATCCCAG
GAATACAAGAAGGCAGAAAAAAGTTTGGAAGATCTTTAAATCTGACAGTGAAAGTGGCTGGT
TACATCCGGCAAGCGGGTGAAGTTCCATCAGGTAATTATTCGAGGTGGAGGACATATTTTACCC
TATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGATTCAATTTATGGAAAAGGATGGGAT
CCTTATGTTGGATAACTACCTTCCCAAAGAGAACATCAGAGGTTTTTCATTGCTGAAAAGAA
AATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAATTATCTTTTCATATCTGCAAGATTTT
TTTCATCAATAAAAAATTATCCTTGAAACAAGTGAGCTTTTGTTTTTTGGGGGGAGATGTTTACT
ACAAAATTAACATGAGTACATGAGTAAGAATTACATTATTTAACTTAAAGGATGAAAGGTATG
GATGATGTGACACTGAGACAAGATGTATAAATGAAATTTTAGGGTCTTGAATAGGAAGTTTAA
ATTTCTTCTAAGAGTAAGTGAAAAGTGCAAGTTGTAACAAACAAAGCTGTAACATCTTTTTCTG
CCAATAACAGAAGTTTGGCATGCCGTGAAGGTGTTTGGAATATTATTGGATAAGAATAGCTC
AATTATCCCAAATAAATGGATGAAGCTATAATAGTTTTTGGGGAAAAGATTCTCAAATGTATAA
AGTCTTAGAACAAAAGAATTCTTTGAAATAAAAAATATTATATATAAAAGTAAAAA

FIGURE 410

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGD SGQPLFLTPYIEAGKIQKGRELSLV
 GPFPGGLNMKSYAGFLTVNKTYSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEHGP
 YVVTSNMTLRDRDFPWTTTSLMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQFFQI
 FPEYKNNDFYVTGESYAGKYVPAIAHLIHS LNPNVREVKINLNGIAIGDGYSDPESIIGGYAEF
 LYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNV TGCSNYY
 NFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVK PWLTEIMNNY
 KVL IYNGQLDIIVAAALTE RSLMGMDWKGSQEYKKA EKVKW KIFKSDSEVAGYIRQAGDFHQV
 IIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
 353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

415/615

FIGURE 411

GCAAGCCAAGGCGCTGTTTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGACATTGTGTACCGCCT
CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA
CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCACCC
CCTGGCCACACTCTTCAAGATCCTGGCGCTCCTTCTACATCAGCCTAGT CATCTTCTACGGCCTCATCTGCATGTA
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG
CGACATCCCCGACGTCAAGAACGACTTCGCCCTTCATGCTGCACCTCATTTGACCAATACGACCCGCTCTACTCAA
GCGCTTCGCGCTCTTCTGTGCGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT
GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCGCC
CAGCATTGCCCAGCTCACGGGCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT
GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA
CGGGCTGCGGGAGCTCAAACGCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCAC
AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCTCAACAGCCTCAA
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCAGCACCT
GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA
CCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGCAAGCTGCG
CTACCTGGACCTCAGCCACAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT
AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG
CAACCGCTGGAGTGCTGCCCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCTGAGCGGAG
GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCCGAGGGGCGAGGCCTAGCTTCTCCCAG
AACTCCCGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGA
GGACAGTATCTGTGGGCTGGCCCCCTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG
CTGCCACCAGAGGTCTGGGACCCTCACTTTAGTTCTTGGTATTTATTTTTCTCCATCTCCACCTCCTTCATCC
AGATAACTTATACATTCCCAAGAAAGTTTCAAGCCAGATGGAAGGTGTTCAAGGAAAGGTGGGCTGCCCTTTTCCCC
TTGTCTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCAG
CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCTCCAGCTGGA
AAGGCCAGGCCTGGAGCTTGCCCTCTCAGTTTTTGTGGCAGTTTTAGTTTTTTGTTTTTTTTTTTAAATCAAA
AAACAATTTTTTTTAAAAAAAAGCTTTGAAAATGGATGGTTTTGGGTATTAAGAAAAAGAAAAAAACTTAAAAAA
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTTCCCTTGAGCAAAGCAGCCAGACGT
TGAAGTGTGTTTCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTCCGGATCTGGTGTGACCTTGGTCCAGGAGTT
CTATTTGTTCTTGGGAGGGAGGTTTTTTTGTGTTTTTTGGGTTTTTTTGGTGTCTTGTGTTTTCTTCTCCTCC
ATGTGTTTGGCAGGCACTCATTTCTGTGGCTGTCGGCCAGAGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG
ACTCGGGTTGGCTAATCCCCGGATGAACGGTGTCTCCATTGCGACCTCCCCCTCCTCGTGCCTGCCCTGCCTCTCCA
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTTCCCCACCTCCTGCGGCATGGGTGTGT
CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTCATGAAGAGCAGACACTTA
GAGGCTGGTCCGGAATGGGAGGTGCCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGC
CTGGAGTGACACAGCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT
AGAAGGGTCCCCGCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAG
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTACAGTATCAAATAAA
ATCTATAACAGAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

416/615

FIGURE 412

MRQTIKVIKFIILICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYIS
LVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSKRF
AVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLEHLFMLSIGIPDTVFDLVELEVKL
ELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRALHIKFTDIKEIPLWIYSLK
TLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSINNEG
KLIVLNSLKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIIISFQHLHRLT
CLKLWYNHAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLTFLPADIG
LLQNLQNLAITANRIETLPPPELFQCRKLRLALHLGNNVLQSLPSRVGELTNLTQIELRGNRLEC
LPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 413

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCATC
TTCATCATTATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**GAGACTCATCAGAA
ACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAGAAG
AAAGGGAAGTATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTGACCC
CAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCAAGATTTT
ATTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGGATCTCA
AAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTGAAGAGTG
TAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTTTGACACCA
TGCCTATCTGTGAGGAAGCTGGCAACATGTACACCTGGAAATCCTAGGTTTGAGTGGGGCAA
AAATACAAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTCTTCTTAGGAT
TCAGAACTCTTCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAACAAAATGCACA
TTGTTTTACCAATGGACACAAATTTCTGGGTCTTTTGCGTGATGGAATCAAGACTTCAAAAA
TATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAAATGCAACGAAATC
TTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGATTTACTCTGGGACG
ACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAACTTTTCAAGATCCGAAATG
TGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTACTCAAATACTGTAATGA
GAACTATAAAATTTGGAGCATGTACATTTTCAAGAGTGTTTTACATTTCAACAGGATAAAATCTATT
TGCTTTTGGACCAAAATGGACATAGAAAACCTGACAAATATCAAATGCACAAATGCCACACATGC
TTTTCCCGAATTATCCTACGAAATTTCCAATATTTAAATTTTGCCAATAATATCTTAACAGACG
AGTTGTTTAAAAGAACTATCCAACCTGCCTCACTTGAAAACCTCTCATTTTGAATGGCAATAAAC
TGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACACCCTTGGAACACTTGGATCTGA
GTCAAAATCTATTACAACATAAAATGATGAAAATTTGCTCATGGCCAGAAACTGTGGTCAATA
TGAATCTGTACATAAATAAATTTGCTGATTCTGTCTTCAGGTGCTTGCCCAAAAGTATTTCAA
TACTTGACCTAAATAAATAACCAATTTCAAACTGTACCTAAAGAGACTATTTCATCTGATGGCCT
TACGAGAACTAAATATTGCATTTAATTTTCTAACTGATCTCCCTGGATGCAGTCATTTTCAAGTA
GACTTTTCAAGTCTGAACATTGAAATGAACCTTCACTCTCAGCCCATCTCTGGATTTTGTTCAGA
GCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAAATCCATTCCGGTGTACCTGTGAATTAA
AAAATTTTCACTCAGCTTGAAACATATTCAGAGGTCATGATGGTTGGATGGTCAGATTTCATACA
CCTGTGAATACCCTTTAAACCTAAGGGGAAGTAAAGAGCGTTTCATCTCCACGAATTTAT
CTTGCAACACAGCTCTGTTGATTGTCAACATTGTGGTTATTATGCTAGTTCTGGGGTTGGCTG
TGGCCTTCTGCTGTCTCCACTTTGATCTGCCCTGGTATCTCAGGATGCTAGGTCAATGCACAC
AAACATGGCACAGGGTTAGGAAAACAACCAAGAACAACTCAAGAGAAATGTCCGATTCCACG
CATTTATTTTACATACAGTGAACATGATTCTCTGTGGGTGAAGAATGAATTGATCCCCAATCTAG
AGAAGGAAGATGGTTCTATCTTGATTGCTTTTATGAAAGCTACTTTGACCCTGGCAAAAGCA
TTAGTGAAAATATTGTAAGCTTCATTGAGAAAAGCTATAAGTCCATCTTTGTTTGTCTCCCA
ACTTTGTCCAGAATGAGTGGTGCCATTATGAATTCTACTTTGCCACCACAATCTCTTCCATG
AAAATCTGATCATATAATTCTTATCTTACTGGAACCCATTCCATTCTATTGCATTCCCACCA
GGTATCATAAACTGAAAGCTCTCCTGGAAAAAAGCATACTTGGAATGGCCCAAGGATAGGC
GTAAATGTGGGCTTTTCTGGGCAAACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCA
GAGAAATGTATGAACTGCAGACATTACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCT
CTCTGATGAGAACAGATTGTCTA**TAA**ATCCACAGTCCTTGGGAAGTTGGGGACCACATACA
CTGTTGGGATGTACATTGATACAACCTTTATGATGGCAATTTGACAATATTTATTAATAAAA
AAATGTTTATTTCCCTTCATATCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACA
CCTTCACAAGTTTATAAGGGCTTATGGAAAAAGGTGTTTCATCCCAGGATTGTTTATAATCATG
AAAAATGTGGCCAGGTGCAGTGGCTCACTCTTGTAAATCCAGCACTATGGGAGGCCAAGGTGG
GTGACCCACGAGGTCAAGAGATGGAGACCATCCTGGCCAACATGGTGAAACCTGTCTCTACT
AAAAATACAAAAATTAGCTGGGCGTGATGGTGACGCTGTAGTCCCAGCTACTTGGGAGGCT
GAGGCAGGAGAATCGCTTGAACCCGGGAGGTGCGAGTTGCAGTGAGCTGAGATCGAGCCACTG
CACTCCAGCCTGGTGACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAA
ATGGAAAACATCCTCATGGCCACAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGT
AATATAATATTACATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTTGGTATGGAAAAA
CATATTAATATGTTATAAACTATTAGGTTGGTGCAAACTAATTGTGGTTTTTGCCATTGAA
TGGCATTGAAATAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGG
GAGGTTGGATTACAGGGAGCATTTGATTCTATGTTGTGTATTTCTATAATGTTTGAATGTT
TAGAATGAATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

418/615

FIGURE 414

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQL
QSSDFHSVSKLRVLILCHNRIQQDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDLSF
NDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLPILN
TTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLLLNKV
DLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFRVFYIQ
QDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNfANNILtDELfKRTIQLPHLKTlI
LNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNdENCsWPETVVNMNLSYNKLSDSVFRCL
PKSIQILDlNNNqIQTVPKETIHLMAReLNIaFNfLTDLPGCSHfSRLSVLNIEMNFILSPS
LDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMVGWSDSYTCEYPLNLRGTRLKDV
HLHELSCNTALLIVTIVVIMLVGLAVAFcCLHFDLPWYLRMLGQCTQTWHRVRKTTQEQlKR
NVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENIVSFIEKSYKSI
FVLSPNFVQNEWCHYEFYFAHhNLFHENSdHIILILLEPIPFYCIPTRYHKLKALLEKKAYLE
WPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLMRTDCL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 415

CGGACGCGTGGGCGGACGCGTGGGCGCTGGGCAAGGGCGGGGCGCCGGGCGGAGCCACCTCTTCCCCCTCCCCCGC
 TTCCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT
 CGGCAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACACAGGCCAGACAGGGGCGCTCG
 CTGCCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGGCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
 AGAAGAGTGGCGGCGGCGGACGGAGAAAACAACCTCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCGGGCTGCCG
 CCGCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTGAGCCCCGGGCCATGGAGCCCCCTGGGGAGGCGG
 CACCAGGGAGCCTGGGCGCCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTTCG
 GCACCTCTGGACAGCCCAGGATGCTGTTGGCCACCCCTCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
 ACCGGATTATTTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTTAGAAGTGCAGGGCACCTTACAGA
 GGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG
 TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCCCTACGCTCCCCCTCTCCAGCCAC
 TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCCGGGGGCAACGTCACCATCACTTACAGCTATG
 CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG
 AGTTTCAGTGCCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
 CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCTCCCTCCCTGCCTTGCAATG
 TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCACCCCCAGT
 CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTG
 GAGATGCAGTGATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
 GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCA
 ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCTGTGGCTTAGGCT
 CTGGCCTGGGAGCTGGCGAAGGCCCTAGGTGAGCGCTGTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
 GTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCTGTGGGGCTGCTGGCACCT
 CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
 GACGCTGTCCGGATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCG
 ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG
 CTGCAGTCATTGGCAGCCTAGTGTGCGGCCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
 TTCGACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
 CTTCCTACGGGCAGCTCATTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTTACAGAGAATCCTAATGATA
 ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
 CCGCCCTCGTCAGCGGGGCGCTTGATGCGACGCCTGGTACGCCGTCTCCGCCGCTGGGGCTTGCTCCCTCGAA
 CCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCAGGTCACACCTTCTGCTGCTCCCCTTGAGGCCCTAGATG
 GTGGCACAGGTCCAGCCCGTGAGGGCGGGGCGAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCATCA
 AGGCTCCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
 TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCC
 CAGGACCAACCCGAGCCCCCTGGACCCACACAGCAGTCTTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
 TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG
 CTCTACTGAGGCCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACTTTTAGAGGTGGGTGAGCCTCCCCTCC
 ACCACTTCCTTCCCTGTCCCTGGATTTCAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT
 TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCCAGTCCCT
 TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA
 GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAGTAGAGA
 CCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGTTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG
 CTCACAAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
 GGAATCATACATCTC

420/615

FIGURE 416

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631
<subunit 1 of 1, 713 aa, 1 stop
<MW: 76193, pI: 5.42, NX(S/T): 4
MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTQLQRPLVRDSRTSPANCTWLILG
SKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAPMG
QGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCGDGVDACGDGSDEAGCSSDPFPGLTTPRPVPS
LPCNVTLEDFYGVFSSPGYTHLASVSHPSCHWLLDPHDGRRRLAVRFTALDLGFGDAVHVYDG
PGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLPWDRPC
GLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGATACYLPA
DRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPDGQDCADGSDEWDCSYVLPKRV
ITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSYGQLIAQGA
IPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGPGARRRQGRMLMRLVRRLLRRWGLLP
RTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKAPLPSASTSPA
PTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPGLGPPGPTRSPPGPHTAVLALEDEDDVLL
VPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 417

GTCGTTCCCTTTGCTCTCTCGCGCCAGTCCCTCCCTGGTTCTCCTCAGCCGCTGTCCGAGGAGAGCACCCGGA
 GACGCGGGCTGCAGTCGCGGCGGCTTCTCCCCGCTGGGCGGCCTCGCCGCTGGGCAGGTGCTGAGCGCCCTAG
 AGCCTCCCTTGCCGCTCCCTCCTCTGCCCCGCGCAGCAGTGACATGGGGTGTGGAGGTAGATGGGCTCCCG
 GCCCCGGAGGCGGCGGTGGATGCGGCGCTGGGCAGAAAGCAGCCGCCGATTCCAGCTGCCCCGCGCGCCCCGGGCG
 CCCCTGCGAGTCCCCGGTTCCAGCCATGGGGACCTCTCCGAGCAGCAGCACCCGCCCTCGCCTCCTGCAGCCGCATC
 GCCCCGCGAGCCACAGCCACGATGATCGCGGGTCCCTTCTCCTGCTTGGATTCCCTTAGCACCACCACAGCTCAG
 CCAGAACAGAAAGCCTCGAATCTCATTTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT
 GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCT
 GTGGGGACCTTTACCAGGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAATG
 ATTGAGAAATTACCTTGTGCTGCCCTTGACTGACCGAGAATGCACCTTGCCACCTGGCATGTTCCAGTCTAACGCT
 ACCTGTGCCCCCATAACGGTGTGTCTGTGGGTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
 TGTAAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT
 CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCCCTTC
 TCCAGCTCCACCTCACCTTCCCTTGGCACAGCCATCTTTCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT
 TCCTCCACTTATGTTCCCAAAGGCATGAACCTCAACAGAAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACTG
 AGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCTC
 CCAAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGAG
 GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG
 CATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCTGCTGCTGGTGTGTTGTGGTGATTGTGGTG
 TGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCGGCAGGATCCAGTGCCATTGTGGAAAAGGCA
 GGGCTGAAGAAATCCATGACTCCAACCCAGAACCAGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
 ATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG
 AGGGAGGTTGTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCTACGCAGCTCTGCAGCACTGGACC
 ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAG
 AAGATTCGTGGGCTGATGGAAGACACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCCG
 CTAGCCCCGAGCCCCATCCCCAGCCCCAACCGGAACTTGAGAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCA
 CAGGACAAGAACAAAGGGCTTCTTCGTGGATGAGTCGGAGCCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC
 TCCGCGCTGAGCAGGAACGGTTCTTTTATTACCAAAGAAAAAGAGACACAGTGTGCGGCAGGTACGCCTGGAC
 CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGAG
 ATTCCCCAGGCTGAGGACAACTAGACCGGCTATTGGAATATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
 CTCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAAATTACTCA
 ATTTAGTGGCAGGGTGGTTTTTAAATTTCTCTGTTTCTGATTTTTGTGTTTGGGGTGTGTGTGTGTGTTGT
 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGAATATGGCCAGTGCTTGAGTTCTTTCTCCTTCTC
 TCTCTCTCTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAACCTGTTGTGAA
 ATACCCACCCTAAAGTTTTTTAAGTTCCATATTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTCTG
 TGCACTTTAAATTTTACTTAACTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC
 TTCTTAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTTATGTCTCTTAACATTACACCTACTTTTT
 AAAACAAATATTACTATTTTATTATTGTTTGTCTTTATAAATTTCTTAAAGATTAAGAAAATTTAAGA
 CCCCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTTCATATTCATGG
 CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTTACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC
 TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT
 GACAACTGGGGCCACCAGAACTTGAACCTTACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACTTT
 GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCAGCTTTGCTTTAAAGATGTCTTG
 TTTTATATACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCTTGGTGGGATTCCTTCACCAATT
 ACTTTAATTAATAATGGCTGCAACTGTAAGAACCCTTGTCTGATATATTTGCAACTATGCTCCCATTTACAAATG
 TACCTTCTAATGCTCAGTTGCCAGGTTCATGCAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTGTGG
 GTAGTGGTGAAGGACCGATATCAGAAAAATGCCCTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTA
 AAAAAAAA

FIGURE 418

```
></usr/seqdb2/ssl/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTSPSSSTALASCSRIARRATATMIAGSLLLLGLFLSTTTAQPEQKASNLIGTYRHVDRATGQ
VLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCPWPMIEKLPCAAL
TDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKKGTTETEDVRCKQCARGTFSDVPSSVMKCKAY
TDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAIFRPEHMETHEVPSSTYVPGKMN
STESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPLNLQVVNHQQGPHHRHILKLLPS
MEATGGEKSSSTPIKGPKRGRHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIIVCSIRKSSRTLK
KGPRQDPSAIVEKAGLKKSMPTPTQNRKWIYYCNGHGIDILKLVAQVGSQWKDIYQFLCNAS
EREVAAFSNGYTADHERAYAAALQHWITIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQL
ETDKLALPMSPLSPSPSPNAKLNSALLTVEPSPQDKNGGFFVDESEPLLRCDSTSSGS
SALSRNGSFITKEKKDVTVLQRQLDPCDLQPIFDDMLHFLNPEELRVIEEIPOAEDKLDRLFE
IIGVKSQEASQTLSDSVYSHLPDLL
```

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

423/615

FIGURE 419

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCCGACTCCACC**ATGA**ACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACA
GAGGCCCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTAC
CAGTTCCTCTGTGGGGCTGGATTTCGGAGGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTTACCCTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAAGTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGATGGGATGGATTATTTGCAG
CAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTTCCAGACTTTATCCTGGAGCCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCAGCCGAGACCAGTGGAGCATGACCCCCAGACAGTGAATGCCTACTACCTTCCAATAAGAATGAGATC
GTCTTCCCCGTGGCATCCTGCAGGCCCCCTTCTATGCCCCGAACACCCCAAGGCCCTGAACTTCGGTGGCATC
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCTGCGG
CACTTCGGCTGCCCTGTGCGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGT**TAG**ACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCCACAGGTGACATGAGTACAGACCCCTCCTCAATCACCACATTG
TGCTCTGCTTTGGGGGTGCCCCCTGCCTCCAGCAGAGCCCCACCATTCACTGTGACATCTTTCCGTGTACCCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 420

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSHS
TCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQAIL
KHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDNFME
VLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTA YLDYM
EELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAPSMDWL
EFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSLDRRFES
AQEKLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMISEIRTAF
EEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSFFQNMLNLY
NFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLP TKNEIVFPAGILQAPFYARNHPKALNFGGI
GVVMGHELTAFDDQGREYDKEGNLRPWWQNESLA AFRNHTACMEEQYNQYQVNGERLNGRQT
LGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCSVRTPESSHEGL
VTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNP GQLCEVW

Type II Transmembrane domain:
amino acids 32-57

FIGURE 421

GGCGCCGCGTAGGCCCGGGAGGCCGGGCCGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGCC
GCCTCTCCGCACG**ATG**TTCCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGCAG
GTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGCCTG
CCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGGCCCG
GGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAGCCGCC
CCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACC GCCTGGCAGTGCTGGTGCCCTT
CCGCGAACGCTTCGAGGAGCTCCTGGTCTTTCGTGCCCCACATGCGCCGCTTCTGAGCAGGAA
GAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAACCGGGCAGC
GCTCATCAACGTGGGCTTCTGAGAGCAGCAACAGCACGGACTACATTGCCATGCACGACGT
TGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCTGAGGCTGGGCCCTTCCACGT
GGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTGCGCGGCATCCTGCTGCT
CTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGGCTGGGGCCGCGA
GGACGACGAGTTCTACCGGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCCGCCCCCTCGGGAAT
CACAAC TGGGTACAAGACATTTCCGCCACCTGCATGACCCAGCCTGGCGGAAGAGGGACCAGAA
GCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGGAGGCCTGAACACTGT
GAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCCCCCTGCACTGTCCTCAA
CATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTCAGCT**TGAG**CTGGATGGAC
AGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGACAAGGCCTCAGGTCGTGG
GCCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCAAGCTACGCAATTGCAGCCA
CCCGGCCGCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGGGTGCTGGGACGCTGCTTGC
CATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCTGTCCGGGACCCCCCTGCCTTCCTGC
TCACCCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGA
CAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGGTAGTGGGGAGGG
CTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

426/615

FIGURE 422

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531
><subunit 1 of 1, 327 aa, 1 stop
><MW: 37406, pI: 9.30, NX(S/T): 1
MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSLGFFSLLWLQLSCSGDVARAVRG
QGQETSGPPRACPPPEPPPEHWEEDASWGPHRLAVLVPPFRERFEELLVFVPHMRRFLSRKKIRH
HIYVLNQVDHFRFNRAALINVGFLESSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVASPE
LHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDFYRRIKGAGLQLFRPSGITTGY
KTFRHLHDPAWRKRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVLNIMLD
CDKTATPWCTFS

Signal peptide:
amino acids 1-42

Transmembrane domain:
amino acids 29-49 (type II)

N-glycosylation site.
amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 27-31

Tyrosine kinase phosphorylation site.
amino acids 226-233

N-myristoylation site.
amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

FIGURE 423

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGATC
TCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATGCT
GCTTTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCCCTC
TCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGTTTTT
GTCACCAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAAGTGGT
GTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCCTGGTGAGGAGCATTAGTAACAG
CTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATGGAGATGG
ATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCCTCCACCAT
CTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTGGAAAAGATTA
TAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTTAGGGCAGGTGGGAAGTCAG
CAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGACTCACCTGGAA
GAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTCATGATCCTCCTTCTTTTTCCTT
TTTCTTCACCTTCATTTAGGCTTTTCTCTGTCTTCATGTCTTGAGATCTCAGAGAATAATA
ATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

428/615

FIGURE 424

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCILLLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWMD
ADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSSTD
VMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

20251010 08:00:00

FIGURE 425

CGGACGCGTGGGCGCCACCTCCGGAACAAGCC**AT**GGTGGCGGCGACGGTGGCAGCGGCGTGG
CTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGGT
AACATCCGGGGCAAACCTGGTGTCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTGAAT
GTGGCCAGCGAGTGCGGCTTACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCGAGAC
CTGGGCCCCCACCACCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGGAGCCT
GACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGACCTACAGTGTCTCATTCCCCATGTTT
AGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCAGACTTCT
GGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGGTGGTAGGG
GCTTGGGACCCAACCTGTGTGAGTGGAGGAGTGCAGACCCAGATCACAGCGCTCGTGAGGAAG
CTCATCCTACTGAAGCGAGAAGACTTAT**TAA**CCACCGCGTCTCCTCCTCCACCACCTCATCCCG
CCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAGAGACCCACTGA
CTCTCCTTCTTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAAAAATTCTAGTAT
TTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAGAGCTCTTGACCAG
TGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAATAGAAGTATATCAA
GCAATAATCTCCACCCAAGGCTTCTGTAACTGGGACCAATGATTACCTCATAGGGCTGTTG
TGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGCCAAATAGGAGGCATTC
AATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAATAAAAACTTGCATCCAAC
ATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTTGTTATTTCTCTGTATTA
TTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCAACAATACCTCACGATATA
AAATAAAAAATGAAAGTATCCTCCTCAAAA

430/615

FIGURE 426

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQH
YRALQQLQRDLGPHHFNVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTGAH
PAFKYLAQTSGKEPTWNFWKYLVPDGGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 427

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGCC
TTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTCTT
CTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGCTTT
GCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCTTCAT
GATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCAAATTT
TTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTTAGGTTT
AAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTATTCTTCACTA
TTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTTTATCCCTG
GGTGCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAATGTTATTCT
AATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTTAAAAGTTTTG
AGTATATTTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

432/615

FIGURE 428

MGVEIAFASVILTCLSLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 429

CCAAAGTGATCATTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCTG
CACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAAATGGTGCTGACCATCTTT
GGGATACAATCTCATGGATACGAGGTTTTTAAACATCATCAGCCCAAGCAACAATGGTGGCAAT
GTTTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGCAGGA
TCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGCTCTCC
CGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAATCTCCAA
TGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTGGGTCAAG
TACAACCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCACCCATTGAG
AAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACATAATGTCGGT
GCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTGTGCAGACATT
CATGTTTAGGATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCCTTGGTTTACAC
TCAAAGTCAAATTAAATTCTTTCCAATGCCCAACTAATTTTGAGATTGAGTCAGAAAATA
TAAATGCTGTATTTATA

434/615

FIGURE 430

><ss.DNA57834
><subunit 1 of 1, 176 aa, 1 stop
><MW: 19616, pI: 7.11, NX(S/T): 0
MVLTIIFIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTTIFDYKH
GYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLESLIK
DVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

Important features:

Signal peptide:

Amino acids 1-26

N-myristoylation sites:

Amino acids 48-54;153-159;156-162;167-173

10055555 011500

435/615

FIGURE 431

GCGTGGGGATGTCTAGGAGCTCGAAGGTGGTGTCTGGGCCTCTCGGTGCTGCTGACGGCGGCCA
CAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTATCA
GAGACATTGAGAGGCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTATTT
TGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAAATCAT
GACTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTGTTGATGGAGA
GTAGCTTAGTAGTATCTTCATCTTTTTTTTTTGGTCACTGTCCTTTTAACTTGATCAAATAAA
GGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGAGTGTGGG
CAGACACTTTTTTGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGGTCACTGTC
CAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGATGGATAGTA
GCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACATCGACCTCAG
AAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

CGTGGGGATG

436/615

FIGURE 432

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILTE
QLEAEREKMLLAKGSQKS

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILTE
QLEAEREKMLLAKGSQKS

FIGURE 433

GAATTTCGTGTCTCGGCACTCACTCCCGGCCGCCCGGACAGGGAGCTTTTCGCTGGCGCGCTTGGCCGGCGACAGGA
 CAGGTTCCGGGACGTCCATCTGTCCATCCGTCCGGAGAGAAATTACAGATCCGCAGCCCCGGGATGCGGGCCGCC
 CGCTGCCGCTGCTGCTGGGCCCTTCTCTCCCGCGCTCTGGCGTAGAGCTATCACTGAGGCAAGGGAAGAAGCCA
 AGCCTTACCCGCTATTCCCGGGACCTTTTCCAGGGAGCCTGCAAACTGACCACACACCGCTGTTATCCCTTCCTC
 ACGCCAGTGGGTACCAGCCTGCCTTGATGTTTTTACCACCCAGCCTGGAAGACCACATACAGGAAACGTAGCCATT
 CCCCAGGTGACCTCTGTGCAATCAAAGCCCCCTACCGCCTCTTGCCCTTCAAACACACAGTTGGACACATAACTTT
 TCTGAACATAAAGGTGTCAAATTTAATTGCTCAATCAATGTACCTAATATATATACCAGGACACCACAATTTCTTGG
 TGGAAAGATGGGAAGGAATTGCTTGGGGGACATCATGAATTACACAGTTTTATCCAGATGATGAAGTTACAGCA
 ATAATCGCTTCTTTCAGCATAACCAGTGTGCAGCGTTCAGACAATGGGTGCTATATCTGTAAGATGAAAAATAAAC
 AATGAAGAGATCGTGTCTGATCCCATCTACATCGAAGTACAAGGACTTCCTCACTTTACTAAGCAGCCTGAGAGC
 ATGAATGTCACCAGAAACACAGCCTTCAACCTCACCTGTGAGGCTGTGGGCCCCGCTGAGCCCCGTCACATTTTC
 TGGGTTCAAACAGTAGCCGTGTTAACGAACAGCCTGAAAAATCCCCCGGCGTGCTAACTGTTCCAGGCCTGACG
 GAGATGGCGGTCTTCAGTTGTGAGGCCCAATGACAAAGGGCTGACCGTGTCCAGGGAGTGCAGATCAACATC
 AAAGCAATTCCTTCCCCACCACTGAAGTCAGCATCCGTAACAGCACTGCACACAGCATTCTGATCTCCTGGGTT
 CCTGGTTTTGATGGATACTCCCCGTTAGGAATTCAGCAGTTCAGGTCAAGGAAGCTGATCCGCTGGGTAATGGC
 TCAGTCATGATTTTTAACACCTCTGCCTTACCACATCTGTACCAATCAAGCAGCTGCAAGCCCTGGCTAATTAC
 AGCATTGGTGTCTTCTGTCATGAATGAAATAGGCTGGTCTGCAGTGAGCCCTTGGATTCTAGCAAGCAGCACTGAA
 GGAGCCCCATCAGTAGCACCTTTAAATGTCACTGTGTTTCTGAATGAATCTAGTGATAATGTGGACATCAGATGG
 ATGAAGCCTCCGACTAAGCAGCAGGATGGAGAATCTGGTGGGCTACCGGATATCCACGTGTGGCAGAGTGCAGGG
 ATTTCCAAAGAGCTCTTGGAGGAAGTTGGCCAGAATGGCAGCCGAGCTCGGATCTCTGTTCAAGTCCACAATGCT
 ACGTGCACAGTGAGGATTGCAGCCGTACCAGAGGGGGAGTTGGGCCCTTCAGTGATCCAGTGAAAAATATTTATC
 CCTGCACACGGTTGGGTAGATTATGCCCCCTCTTCAACTCCGGCGCCTGGCAACGCAGATCCTGTGCTCATCATC
 TTTGGCTGCTTTTGTGGATTTATTTTGATTGGGTTGATTTTATACATCTCCTTGGCCATCAGAAAAAGAGTCCAG
 GAGACAAAGTTTGGGAATGCATTCACAGAGGAGGATTCTGAATTAGTGGTGAATTATATAGCAAAGAAATCCTTC
 TGTCGGCGAGCCATTGAACTTACCTTACATAGCTTGGGAGTCAGTGAGGAACACAAAAATAAACTAGAAGATGTT
 GTGATTGACAGGAATCTTCTAATTCTTGGAAAAATCTGGGTGAAGGAGAGTTTGGGTCTGTAATGGAAGGAAAT
 CTTAAGCAGGAAGATGGGACCTCTCTGAAAGTGGCAGTGAAGACCATGAAGTTGGACAACCTCTTCACATCGGGAG
 ATCGAGGAGTTTCTCAGTGAGGCAGCGTGCATGAAAGACTTCAGCCACCCAAATGTCATTGACTTCTAGGTGTG
 TGTATAGAAATGAGCTCTCAAGGCATCCCCAAAGCCCATGGTAATTTTACCCTTCATGAAATACGGGGACCTGCAT
 ACTTACTTACTTTATTTCCCGATTGGAGACAGGACCAAGCATATTCCTCTGCAGACACTATTGAAGTTCATGGTG
 GATTTGCCCTGGGAATGGAGTATCTGAGCAACAGGAATTTTCTTCATCGAGATTTAGCTGCTCGAAACTGCATG
 TTGCGAGATGACATGACTGTCTGTGTTGCGGACTTCGGCCTCTCTAAGAAGATTTACAGTGGCGATTATTACCGC
 CAAGGCCGATTTGCTAAGATGCCGTGTTAAATGGATCGCCATAGAAAGTCTTGACAGCCGAGTCTACACAAGTAAA
 AGTGATGTGTGGGCATTTGGCGTGACCATGTGGGAAATACGTACGCGGGGAATGACTCCCTATCCTGGGGTCCAG
 AACCATGAGATGTATGACTATCTTCTCCATGGCCACAGGTTGAAGCAGCCCCGAAGACTGCCTGGATGAACTGTAT
 GAAATAATGTACTCTTGCTGGAGAACCGATCCCTTAGACCGCCCCACCTTTTTCAGTATTGAGGCTGCAGCTAGAA
 AAACCTCTTAGAAAGTTTGCTGACGTTTCGGAACCAAGCAGACGTTATTTACGTCAATACACAGTTGCTGGAGAGC
 TCTGAGGGCCTGGCCCAGGGCCCCACCTTGTCTCCACTGGACTTGAACATCGACCCTGACTCTATAATTGCCCTCC
 TGCACCTCCCCGCGCTGCCATCAGTGTGGTTCACAGCAGAAGTTCATGACAGCAAACCTCATGAAGGACGGTACATC
 CTGAATGGGGGCAGTGAGGAATGGGAAGATCTGACTTCTGCCCCCTCTGCTGCAGTCAAGCTGAAAAGAACAGT
 GTTTTACCGGGGGAGAGACTTGTTAGGAATGGGGTCTCCTGGTCCCATTCGAGCATGCTGCCCTTGGGAAGCTCA
 TTGCCCGATGAACTTTTGTGTTGCTGACGACTCCTCAGAAGGCTCAGAAGTCTTGATGTAGGAGAGGTGCGGGGA
 GACATTTCCAAAAATCAAGCCAATTTCTTCTGCTGTAGGAGAATCCAATTGTACCTGATGTTTTTGGTATTTGTCTT
 CCTTACCAAGTGAACCTCATGGCCCCAAAGCACCAGATGAATGTTGTTAAGGAAGCTGTCAATAAAAATACATAA
 TATATATTTATTTAAAGAGAAAAATATGTGTATATCATGAAAAAGACAAGGATATTTTAATAAAACATTACTTA
 TTTTATTTTCACTTATCTTGCATATCTTAAATTAAGCTTCAGCTGCTCCTTGATATTAACCTTTGTACAGAGTTG
 AAGTTGTTTTTTCACTTCTTTTCTTTTTCATTACTATTAAATGTAAAAATATTTGTAAAAATGAAATGCCATATTT
 TGACTTGGCTTCTGGTCTTGATGTATTTGATAAGAATGATTAATTTTCTGATATGGCTTCATAATAAAATTGAA
 ATAGGA

438/615

FIGURE 434

MGPAPLPLLLGLFLPALWRRRAITEAREEAKPYPLFPFPFPGSLQTDHTPLLSLPHASGYQPALMFSPTQPGRPHT
GNVAIPQVTSVESKPLPPLAFKHTVGHIILSEHKGVKFNC SINVPNIYQDTTISWWKDGKELLGGHHRITQFYPD
DEVTAIIASF SITSVQRSDNGSYICKMKINNEEIVSDPIYIEVQGLPHFTKQPESMNVTRNTAFNLTCQAVGPPE
PVNIFWVQNSSRVNEQPEKSPGVLTVPGLTEMAVFSCEAHNDKGLTVSQGVQINIKAIKPSPTTEVSIRNSTAHSI
LISWVPGFDGYSPFRNC SIQVKEADPLGNGSVMIFNTSALPHLYQIKQLQALANYSIGVSCMNEIGWSAVSPWIL
ASTTEGAPSVAPLNVTVFLNESSDNVDIRWMKPPTKQQDGELVGYRISHVWQSAGISKELLEEVGQNGSRARISV
QVHNATCTVRIA AVTRGGVGPFS DPVKIFIPAHGWVDYAPSS TPAPGNADPVLIIIFGCFCGFILIGLILYISLAI
RKRVQETKFGNAFTEEDSELVVNYIAKKSFCRRAIELTLHSLGVSEELQNKLEDVVIDRNLILGKILGEGEFGS
VMEGNLKQEDGTSLKVAVKTMKLDNSSHREIEEFLSEAACMKDFSHPNVIRLLGVCIEMSSQGI PKPMVILPFMK
YGD LHTYLLYSRLETGPKHIPLQTL LKFMVDIALGMEYLSNRNFLHRDLAARNCMLRDDMTVCVADFGLSKKIYS
GDYYRQGRIAKMPVKWIAIESLADRVYTSKSDVWAFGVTMWEIRTRGMTYPYGVQNHMYDYLLHGHRLKQPEDC
LDELYEIMYSCWR TDPLDRPTFSVLRLQLEKLLLESLPDVRNQADVIVNTQLLESSEGLAQGPTLAPLDLNI DP
SIIASCTPRAAISVVTAEVHDSKPHEGRYILNGGSEEWEDLTSA PSAAVTAEKNSVLPGERLVRNGVSWSHSSML
PLGSSLPDELLFADDSSEGSEVLM

Signal sequence:

Amino acids 1-18

Transmembrane domain:

Amino acids 501-520

N-glycosylation sites:

Amino acids 114-118;170-174;207-211;
215-219;234-238;294-298;316-320;329-333;
336-340;354-358;389-393;395-399;442-446;
454-458;625-629

Tyrosine kinase phosphorylation sites:

Amino acids 675-683;865-873;923-930

N-myristoylation sites:

Amino acids 41-47;110-116;171-177;
269-275;275-281;440-446;507-513;535-541;
966-972

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 351-362

Tyrosine protein kinases specific active-site signature:

Amino acids 719-732

FIGURE 435

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGATG
 TGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTAGTGTCCGATTCTGATTCCGGCAAGGAT
 CCAAGC**ATG**GAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTCCTG
 CTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTGGGGC
 CCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCTGCCTG
 AGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGACTGCCCA
 CCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCACCATGGC
 CAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCAAGTGCCAA
 GCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACGCGTTGCTAT
 ACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGATCACCAGCTG
 GGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCACCTGCCGGCTG
 GTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACTGTGGTTGCACTT
 CCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAGGTCTCTGATCACTTATATCTGGAAACC
 AAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAACCTTTCCTTGTGGAC
 AATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGAATGGCTGGACCACTC
 ACAGCAGATTTTATTGTCAAGATTCGTAACCTCGGGCTCCGCTGACAGTACAGTCCAGTTCATC
 TTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGT
 GGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCT
 GACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAAC
 TTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTTACCATCCC
 CTTCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGGCATCCAG
 AGCCGGGCAGTTTCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAGAGTGG
 AAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTTGACTGCCCTAAA
 TGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTCAGATACCGTGTG
 GTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCCAAAAACAAAGCCCCAC
 ATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAACCCAAAGAGAAACTTCCAGTCGAG
 GCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGTGTCTCAGAGGAG
 CCCTCG**TAA**AGTTGTAAAAGCACAGACTGTTCTATATTTGAAACTGTTTTGTTTAAAGAAAGCA
 GTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGAACTAAGTGTAATCATCTCACCAAAGCTTT
 TTGGCTCTCAAATTAAAGATTGATTAGTTTCAAAAAAAAAA

440/615

FIGURE 436

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLSS
KSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPSLKCQAK
GTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVR
GQYKSQLSATKSDDTVVALPYGSRHIRLVKGPDLHLYLETKTLQGTKGENSLSSSTGTFLVDNS
SVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFPCSATCGG
GYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLP
RWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL
AQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPVEAK
LPWFKQAQEELEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

2025-10-10 09:55:00

FIGURE 437

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATGT**GGTTACCCCTTGGTCTCCTGT
CTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCAGC
AGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATGTTA
TGGATGTTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGTGATC
TCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCT**TGAA**AATCTGGCATGAGATG
GCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCTCCTCCCAAGTCTGT
TCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAGCAATCTG
TGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGTAGCCACCT
CCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTACAGGACATTCTGC
CCCTGTGTGCCACCAAACCAGGACTTTCCTTGGCTTGGCATCCCTGGCTCTCTCCTGGTAC
CCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACTATGGCGATGGC
CATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCAGAGGGGAAATGGG
GCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAAGGGAAGCAACAGG
AACTTCTGCAACTGGTTTTTTATCGGAAAGATCATCCTGCCTGCAGATGCTGTTGAAGGGGCAC
AAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAAATAGAACAGTCTGCT
GGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGGAAGTCACTCAGCCTCC
CCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGGAACCAAGGAAACTAACA
ATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCCCATGCCACCAAAACAATA
AAACAAAATTCTCTAACACTGAAA

442/615

FIGURE 438

MWLPLGLLSLCLSPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQSGMEH
RNHLCFCDLYDRATSPPLKCSLL

[illegible]

FIGURE 439

GTTCTCATAGTTGGCGTCTTCTAAAGGAAAAAACTAAAAATGAGGAACTCAGCGGACCGGGAGCGACGCAGCTT
 GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAGCCGAGTGGCCCGAGGTGTCTGAGGGGCTGG
 GGCAAAGGTGAAAGAGTTTTCAGAACAGCTTCTTGGAACCCATGACCCATGAAGTCTTGTGACATTTTATACCGT
 CTGAGGGTAGCAGCTCGAACTAGAAAGTGGAGTGTGTCAGGGACGGCAGTATCTCTTGTGTGACCCCTGGC
 GGCCTATGGGACGTTGGCTTCAGACCTTTGTGATACACC**ATG**CTGCGTGGGACGATGACGGCGTGAGAGGAAATG
 AGGCCTGAGGTCACTGGCTTGCCTCCTCCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG
 GTCACCGTCCAGCCTGCCGTAACCGTCCAGAACCCCGGAGGCACTGTGATCTTGGGCTGCGTGAGAACCTCCA
 AGGATGAATGTAACCTGGCGCCTGAATGGAAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGTCTCATCACC
 CACGGGACCCCTCGTCATCACTGCCCTTAACAACCACTGTGGGACGGTACCAGTGTGTGGCCCGGATGCCCTGCC
 GGGGCTGTGGCCAGCGTGCCAGCCACTGTGACACTAGCCAACTCCAGGACTTCAAGTTAGATGTGCAGCACGTG
 ATTGAAGTGGATGAGGGAAACACAGCAGTCACTTGCCTGCCACCTGCCCTGAGAGCCACCCCAAAGCCAGGTCCGG
 TACAGCGTCAAACAAGAGTGGCTGGAGGCCTCCAGAGGTAACCTGATCATGCCCTCAGGGAACCTCCAGATT
 GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGCAGCCTACAACCCAGTGACCCAGGAAGTGAACCC
 TCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCGCTGAGGCTGCCCGCATCATCTACCCCCAGGAGCC
 CAAACCATCATCGTACCCAAAGGCCAGAGTCTCAATCTGGAGTGTGTGGCCAGTGGAAATCCACCCCCAGGGTCT
 ACCTGGGCCAAGGATGGGTCCAGTGTACCCGGCTACAACAAGACGCGCTTCTGCTGAGCAACCTCCTCATCGAC
 ACCACAGCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGGTGGGGCAGCCCGGGCAGCGGTCT
 ATCCTCTACAATGTTCCAGTGTGTTGAACCCCTGAGGTACCATGGAGCTATCCAGTGGTTCATCCCTGGGGC
 CAGAGTGGCAAGCTTACCTGTGAGGTGCGTGGGAACCCCGCCCTCCGTGCTGTGGCTGAGGAATGCTGTGCCC
 CTCTATCTCCAGCCAGCGCCTCCGGCTCTCCCGCAGGGCCCTGCGCGTGTCTCAGCATGGGGCTGAGGACGAAGGC
 GTCTACCAAGTGCATGGCCGAGAACGAGGTTGGGAGCGCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC
 ATAACCCCAAGGCTATGGCAGGATGCTGAGCTGGCTACTGGCACACCTCCTGTATCACCTCCAACTCGGCAAC
 CCTGAGCAGATGCTGAGGGGGCAACCGGCGCTCCCCAGACCCCAACGTCACTGGGGCTGCTTCCCGAAGTGT
 CCAGGAGAGAAGGGGCGAGGGGCTCCCGCCGAGGCTCCCATCATCTCAGCTCGCCCCGCACCTCCAAGACAGAC
 TCATATGAAGTGGTGTGGCGGCCTCGGCATGAGGGCAGTGGCCGGGCGCAATCCTCTACTATGTGGTGAACAC
 CGCAAGCAGGTCAAAATTCCTCTGACGATTGGACCATCTCTGGCATTCCAGCCAACGAGACCCGCTGACCTC
 ACCAGACTTGACCCCGGGAGCTTGTATGAAGTGGAGATGGCAGCTTACAAGTGTGCGGGAGAGGGCCAGACAGCC
 ATGGTCACCTTCCGAAGTGGACGGCGGCCCAACCCGAGATCATGGCCAGCAAAGAGCAGCAGATCCAGAGAGAC
 GACCCCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCAGGCGCGCTCTCCCCCAGAAAGTCCCCGACAGG
 CCCACCATCTCCACGGCCTCCGAGACCTCAGTGTACGTGACCTGGATTCCCGGTGGGAATGGTGGGTCCCAATC
 CAGTCTTCCGTGTGGAGTACAAGAAGCTAAAGAAAGTGGGAGACTGGGATTCTGGCCACCAGCGCCATCCCCCA
 TCGCGGCTGTCCGTGGAGATCACGGGCTTAGAGAAAGGCACCTCCTACAAGTTTCGAGTCCGGGCTCTGAACATG
 CTGGGGGAGAGCGAGCCAGCGCCCCCTCTCGGCCCTACGTGGTGTGCGGCTACAGCGGTGCGGTGTACGAGAGG
 CCCGTGGCAGGTCTTATATACCTTACGGATGCGGTCAATGAGACCACCATCATGCTCAAGTGGATGTACATG
 CCAGCAAGTAAACAACACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTGACAATGATAGT
 GACTACAAGAAGGATATGGTGAAGGGGACAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGAGACCTCCTAC
 GACATTAAGATGCAGTGTCTCAATGAAGGAGGGGAGAGCGAGTTCAGCAACGTGATGATCTGTGAGACCAAAGCT
 CGGAAGTCTTCTGGCCAGCCTGGTGCAGTGCACCCCAACTCTGGCCCCACCACAGCCGCCCCCTTCTGAAACC
 ATAGAGCGGCGGTGGGCACTGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTGCGGGTCTGTC
 CTGGGCTCCATCGTTCTCATCATCGTCACCTTCATCCCCTTCTGCTTGTGGAGGGCCTGGTCTAAGCAAACCAT
 ACAACAGACCTGGGTTTTCTCGAAGTGGCCTTCCACCCTCCTGCCCCGTATACTATGGTGGCATTGGGAGGACTC
 CCAGGCCACCAGGCCAGTGGACAGCCCTACCTCAGTGGCATCAGTGGACGGGCTGTGCTAATGGGATCCACATG
 AATAGGGGCTGCCCCCTCGGCTGCAGTGGGCTACCCGGGCATGAAGCCCCAGCAGCACTGCCAGGCGAGCTTCAG
 CAGCAGAGTGACACCAGCAGCCTGCTGAGGCAGACCCATCTTGGCAATGGATATGACCCCAAAGTACCAGATC
 ACGAGGGGTCCCAAGTCTAGCCCGGACGAGGGCTCTTTCTTATACACACTGCCCGACGACTCCACTCACCAGCTG
 CTGCAGCCCCATCACGATGCTGCCAACGCCAGGAGCAGCCTGTGCTGTGGGCCAGTCAGGGGTGAGGAGAGCC
 CCGACAGTCTGTCTGGAAGCAGTGTGGGACCCCTCCATTTCACTCAGGGCCCCCATGCTGCTTGGGCCCTTGTG
 CCAGTTGAAGAGGTGGACAGTCTGACTCCTGCCAAGTGAGTGGAGGAGACTGGTGTCCCCAGCACCCCGTAGGG
 GCCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCCGGGGCCACTGGTGGTGTGCTTTTGAACACCCACCT
 CTCACAATTT**TAG**GCAGAAGCTGATATCCAGAAAGACTATATATTTTATTTTAAAAAAGAGAAAA
 AGAGACAGAGAAAATTGGTATTTATTTTCTATTATAGCCATATTTATATATTTATGCACCTGTAAATAAATGTA
 TATGTTTTATAATTCTGGAGAGACATAAGGAGTCTACCCGTTGAGGTTGGAGAGGGAAAAATAAGAAGCTGCCA
 CCTAACAGGAGTCAACCCAGGAAAGCACCGCACAGGCTGGCGCGGGACAGACTCCTAACCTGGGGCCTCTGCAGTG
 GCAGGCGAGGCTGCAGGAGGCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTTCATCAGGACCA
 TGAGGGAACAGCAAGGGGCACGGTATCACAGCCTGGAGACACCCACACAGATGGCTGGATCCGGTGTCTACGGGAA
 ACATTTTCTTAAGATGCCCATGAGAACAGACCAAGATGTGTACAGCACTATGAGCATTAAAAAACCTTCCAGAA
 CAATAATCCGTGGCAACATATCTCTGTAAAAACAAACACTGTAACTTCTAAATAAATGTTTAGTCTTCCCTGTAAAA

FIGURE 440

MLRGTMTAWRGMRPEVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPPR
MNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTLAN
LQDFKLDVQHVIEWDEGNTAVIACHLPESHHPKAQVRYSVKQEWLEASRGNYLIMPSGNLQIVN
ASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARI IYPPEAQTIIIVTKGQSLILEC
VASGI PPPRV TWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGAAVILY
NVQVFEPPEVTMELSQLVIPWGQSAKLTCEVRGNPPPSVLWLRLNAVPLISSQRLRLSRRALRV
LSMGPEDEGVYQCAENEVGSAAHVQLRTSRPSITPRLWQDAELATGTPPVSPSKLGNPEQM
LRGQPALPRPPTSVGPASPKCPGEGKGQGAPEAPI ILSSPRTSKTDSYELVWRPRHEGSGRAP
ILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVEMAAYNCAGEGQTAMVTFR
TGRRPKPEIMASKEQQIQRDDPGAS PQSSSQPDHGRLSPPEAPDRPTISTASETSVYVTWIPR
GNGGFPIQSF RVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGT SYKFRVRALNMLGESEP
SAPSRPYVVS GYSGRVYERPVAGPYITFTDAVNETTIMLKWMIYPASNNNTPIHGFIYYRPT
DSDNDS DYKKDMVEGDKYWH SISHLQPETS YDIKMQCFNEGGESEFSNVMICETKARKSSGQP
GRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDL PYLIVGVVLGSI VLIIVTFIPFCLWRAW
SKQKHTTDLGFPR SALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRACANGIHMNRGCPSAA
VGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPKSSPDEGSFLYTL PDDS
THQLLQPHHDCCQRQE QPAAVGQSGVRRAPDSPVLEAVWDPPFHSGPPCCLGLVPVEEVDSPD
SCQVSGGDWCPQHVPV GAYVGQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 16-30 (type II), 854-879

FIGURE 441

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGGC
TCTGGCAGGCTCCTGGCAGC**ATGG**CAGTGAAGCTTGGGACCCCTCCTGCTGGCCCTTGCCCTGG
GCCTGGCCCCAGCCAGCCTCTGCCCCGCCGAAGCTGCTGGTGTTCCTGCTGGATGGTTTTCGCT
CAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTTTTCAAAGAGATTGTGAGCAGGG
GAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTATTATACCC
TAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACCTACATGTGGGACCCACCA
CCAACAAGTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTGGAATGGAT
CAGAACCTCTGTGGGTCACTCTGACCAAGGCCAAAAGGAAGGTCTACATGTACTACTGGCCAG
GCTGTGAGGTTGAGATTCTGGGTGTGAGACCCACCTACTGCCTAGAATATAAAAATGTCCCAA
CGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAAGAGTGGCCGGGCCG
ACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGGGCCTGCATCTC
CGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGACCAAGTGGATCC
AGGAGCGGGGCCTGCAGGACCGCCTGAACGTCATTATTTTCTCGGATCACGGAATGACCGACA
TTTTCTGGATGGACAAAGTGATTGAGCTGAATAAGTACATCAGCCTGAATGACCTGCAGCAAG
TGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCCTGGGAAACACTCTGAGATATATA
ACAACTGAGCACAGTGAACACATGACTGTCTACGAGAAAGAAGCCATCCCAAGCAGGTTCT
ATTACAAGAAAGGAAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGATGAAGGCTGGTTCATAA
CTGAGAATCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCAGGCGGGAAGGTTGGCAGC
GTGGATGGCACGGCTACGACAACGAGCTCATGGACATGCGGGGCATCTTCCTGGCCTTCGGAC
CTGATTTCAAATCCAACCTTCAGAGCTGCTCCTATCAGGTGCGGTGGACGTCTACAATGTCATGT
GCAATGTGGTGGGCATCACCCCGCTGCCCAACAACGGATCCTGGTCCAGGGTGATGTGCATGC
TGAAGGGCCGCGCCGGCACTGCCCCGCCTGTCTGGCCCAGCCACTGTGCCCTGGCACTGATTC
TTCTCTTCCTGCTTGCA**TAA**CTGATCATATTGCTTGTCTCAGAAAAAACACCATCAGCAAAG
TGGGCCTCCAAAGCCAGATGATTTTCATTTTATGTGTGAATAATAGCTTCATTAACACAATCA
AGACCATGCACATTGTAAATACATTATTCTTGATAATTCTATACATAAAAGTTCCTACTTGT
TAAA

446/615

FIGURE 442

MAVKLGTLTLLALALGLAQPASARRKLLVFLLDGFRSDYISDEALES LPGFKEIVSRGVKVDYL
TPDFPSLSYPNYYTLMTGRHCEVHQMIGNYMWDPTTNKSFDIGVNKDSL MPLWNGSEPLWVT
LTKAKRKVYMYYPGCEVEILGVRPTYCLEYKNVPTDINFANAVSDALDSFKSGRADLAAIYH
ERIDVEGHHYGPASPQRKDALKAVDTVLKYMTKWIQERGLQDRLNVIIFSDHGMTDIFWMDKV
IELNKYISLNDLQQVKDRGPVVS LWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRFYKKGKF
VSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGIFLAFGPDFKSNF
RAAPIRSVDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALILLFLLA

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 100-104, 118-122, 341-345, 404-408

N-myristoylation sites.

amino acids 148-154, 365-371

Amidation site.

amino acids 343-347

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 443

AGTGACTGCAGCCTTCCTAGATCCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAGC
ACCAGTGTGTGAGGGGAGCAGGCAGCGGTCTAGCCAGTTCCTTGATCCTGCCAGACCACCCA
GCCCCCGGCACAGAGCTGCTCCACAGGCACCAATGAGGATCATGCTGCTATTACAGCCATCCT
GGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGGTTCC
TGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTCAAAAG
CCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAAGGAATC
AACATCTCCCAGAAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGAGGAGCGT
CCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCCCCTTCATCC
CAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACCTTTATAAGA
CTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGTATCCCCCGAG
AGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTCTTTCCCTGTCC
CAATCCCCAGGTGCGCACGCTCCTGTTACCCCTTCTCTTCCCTGTTCTTGTAACATTCTTGTG
CTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAACTGCATAGTGAATATCCC
CAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGTCTACATTAAAAAT
ATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTTGCATATGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

448/615

FIGURE 444

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLKA
LSQASTDPKESTSPKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGKSS
LGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

449/615

FIGURE 445

TGGACTTCTCTGGACCACAGTCCTCTGCCAGACCCCTGCCAGACCCAGTCCACCATGATCCATCTGGGTCACAT
CCTCTTCCTGCTTTTGCTCCCAGTGGCTGCAGCTCAGACGACTCCAGGAGAGAGATCATCACTCCCTGCCTTTTA
CCCTGGCACTTCAGGCTCTTGTTCCGGATGTGGGTCCCTCTCTCTGCCGCTCCTGGCAGGCCTCGTGGCTGCTGA
TGCGGTGGCATCGCTGCTCATCGTGGGGGCGGTGTTCTGTGCGCACGCCCACGCCGCAGCCCCGCCAAGATGG
CAAAGTCTACATCAACATGCCAGGCAGGGGCTGACCCTCCTGCAGCTTGGACCTTTGACTTCTGACCCCTCTCATC
CTGGATGGTGTGTGGTGGCACAGGAACCCCGCCCCAACTTTTGGATTGTAATAAAACAATTGAAACACCA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615

450/615

FIGURE 446

MIHLGHILFLLLLPVAAAQTPGERSSLPAFYPGTSGSCSGCSLSLPLLAGLVAADAVASLLIVGAVFLCARPR
RSPAQDGKVYINMPGRG

Signal peptide:	Amino acids 1-18
transmembrane domain:	Amino acids 51-70
Glycosaminoglycan attachment site:	Amino acids 40-44
N-myristoylation sites:	Amino acids 34-40;37-43;52-58
Prokaryotic membrane lipoprotein lipid attachment site:	Amino acids 29-40

2025-10-27 10:56:50

451/615

FIGURE 447

GCCAGGTGTGCAGGCCGCTCCAAGCCCAGCCTGCCCCGCTGCCGCCACC**ATG**ACGCTCCTCCC
CGGCCTCCTGTTTCTGACCTGGCTGCACACATGCCTGGCCCCACCATGACCCCTCCCTCAGGGG
GCACCCCCACAGTCACGGTACCCACACTGCTACTCGGCTGAGGAACTGCCCCCTCGGCCAGGC
CCCCCACACCTGCTGGCTCGAGGTGCCAAGTGGGGGCAGGCTTTGCCTGTAGCCCTGGTGTC
CAGCCTGGAGGCAGCAAGCCACAGGGGGAGGCACGAGAGGCCCTCAGCTACGACCCAGTGCCC
GGTGCTGCGGCCGGAGGAGGTGTTGGAGGCAGACACCCACCAGCGCTCCATCTCACCCCTGGAG
ATACCGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGTGCCCTGTG
CAGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTGCGGGCTGCT
CCAGAGCCTGCTGGTGCTGCGCCGCGGCCCTGCTCCCGCGACGGCTCGGGGCTCCCCACACC
TGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTGCGCTGCACCTGCGTGCTGCC
CCGTTTCAGT**GTA**CCGCCGAGGCCGTGGGGCCCCCTAGACTGGACACGTGTGCTCCCCAGAGGG
CACCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCAACACTACCCTTGGGGTC
TGGGCATTCCCCGTGTCTGGAGGACAGCCCCCACTGTTCTCCTCATCTCCAGCCTCAGTAGT
TGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTGCAGAAAAGGTGTCACACGG
CTGCCTGTACCTTGGCTCCCTGTCTGCTCCCGGCTTCCCTTACCCTATCACTGGCCTCAGGC
CCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCTTAAACAATTATTTAAG
TGTACGTGTATTATTAAACTGATGAACACATCCCCAAAA

2025.10.09 09:52:00

452/615

FIGURE 448

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHELLARGAKWGQALP
VALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTDED RYPQKLAF
AECLCRGCIDARTGRETAALNSVRLQLSLLVLRRRRPCSRDGSGLPTPGAFAFHTEFIHVPVGC
TCVLPRSV

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 112-121

N-myristoylation sites.

amino acids 32-38, 55-61, 133-139

Leucine zipper pattern.

amino acids 3-25

Homologous region to IL-17.

amino acids 99-195

2025-03-04 10:00:00

FIGURE 449

TGCAGAGCTTGTGGAGGCC**ATG**GGGCGCGTCGTGCGGAGCTCGTCTCCTCGCTGCTGGGGTT
GTGGCTGTTGCTGTGCAGCTGCGGATGCCCCGAGGGCGCCGAGCTGCGTGCTCCGCCAGATAA
AATCGCGATTATTGGAGCCGGAATTGGTGGCACTTCAGCAGCCTATTACCTGCGGCAGAAATT
TGGGAAAGATGTGAAGATAGACCTGTTTGAAAGAGAAGAGGTCGGGGGGCCGCCTGGCTACCAT
GATGGTGCAGGGGCAAGAATACGAGGCAGGAGGTTCTGTTCATCCATCCTTTAAATCTGCACAT
GAAACGTTTTGTCAAAGACCTGGGTCTCTCTGCTGTTTCAGGCCTCTGGTGGCCTACTGGGGAT
ATATAATGGAGAGACTCTGGTATTTGAGGAGAGCAACTGGTTCATAATTAACGTGATTAAATT
AGTTTGGCGCTATGGATTTCAATCCCTCCGTATGCACATGTGGGTAGAGGACGTGTTAGACAA
GTTTCATGAGGATCTACCGCTACCAGTCTCATGACTATGCCTTCAGTAGTGTGAAAAATTACT
TCATGCTCTAGGAGGAGATGACTTCCTTGGAATGCTTAATCGAACACTTCCTTGAAACCTTGCA
AAAGGCCGGCTTTTCTGAGAAGTTCCTCAATGAAATGATTGCTCCTGTTATGAGGGTCAATTA
TGGCCAAAGCACGGACATCAATGCCTTTGTGGGGCGGTGTCACTGTCCTGTTCTGATTCTGG
CCTTTGGGCAGTAGAAGGTGGCAATAAACTTGTTTGCTCAGGGCTTCTGCAGGCATCCAAAAG
CAATCTTATATCTGGCTCAGTAATGTACATCGAGGAGAAAAACAAAGACCAAGTACACAGGAAA
TCCAACAAAGATGTATGAAAGTGGTCTACCAAATTGGAAGTGAAGTTCAGACTTCTATGA
CATCGTCTTGGTGGCCACTCCGTTGAATCGAAAAATGTGGAATATTACTTTTCTCAACTTTGA
TCCTCCAATTGAGGAATTCCATCAATATTATCAACATATAGTGACAACTTTAGTTAAGGGGGA
ATTGAATACATCTATCTTTAGCTCTAGACCCATAGATAAAATTTGGCCTTAATACAGTTTTAAC
CACTGATAATTGAGATTTGTTTCAATTAACAGTATTGGGATTGTGCCCTCTGTGAGAGAAAAGGA
AGATCCTGAGCCATCAACAGATGGAACATATGTTTGGAAGATCTTTTCCAAGAAACTCTTAC
TAAAGCACAAATTTTAAAGCTCTTTCTGTCTTATGATTATGCTGTGAAGAAGCCATGGCTTGC
ATATCCTCACTATAAGCCCCCGGAGAAATGCCCTCTATCATTCTCCATGATCGACTTTATTA
CCTCAATGGCATAGAGTGTGCAGCAAGTGCCATGGAGATGAGTGCCATTGCAGCCCACAACGC
TGCACTCCTTGCCTATCACCGCTGGAACGGGCACACAGACATGATTGATCAGGATGGCTTATA
TGAGAACTTAAAACTGAACTA**TGA**AGTGACACACTCCTTTTTCCCCTCCTAGTTCCAAATGA
CTATCAGTGGCAAAAAAGAACAAAATCTGAGCAGAGATGATTTTGAACCAGATATTTTGCCAT
TATCATTGTTTAAATAAAAGTAATCCCTGCTGGTCATAGGAAAAAAAAAAAAA

454/615

FIGURE 450

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62880

<subunit 1 of 1, 505 aa, 1 stop

<MW: 56640, pI: 6.10, NX(S/T): 4

MGRVVAELVSSLLGLWLLLCSCGCPEGAELRAPDKIAIIGAGIGGTSAAYYLRQKFGKDVKI
DLFEREEVGGRLATMMVQGGQYEAGGSVIHPLNLHMKRFVKDLGLSAVQASGLLGIYNGETL
VFEEASNWFIINVIKLVWRYGFGQSLRMHMMWVEDVLDKFMRIYRYQSHDYAFSSVEKLLHALGGD
DFLGMLNRTLLETQLQAGFSEKFLNEMIAPVMRVNYGQSTDINAFVGAVSLSCSDSGLWAVEG
GNKLVCSGLLQASKSNLISGSVMYIEEKTCTKYTGNTKMYEVVYQIGTETRSDFYDIVLVAT
PLNRKMSNITFLNFDPPIEEFHQYYQHIVTTLVKGELNTSIFSSRPIDKFGLNTVLTTDNSDL
FINSIGIVPSVREKEDPEPSTDGTYVWKIFSQETLTKAQILKLFLSYDYAVKKPWLAYPHYKP
PEKCPSIILHDLRYLNGIECAASAMEMSAIAAHNAALLAYHRWNGHTDMIDQDGLYEKLKTEL

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation sites.

amino acids 196-200, 323-327, 353-357

Tyrosine kinase phosphorylation site.

amino acids 291-298

N-myristoylation sites.

amino acids 23-29, 41-47, 43-49, 45-51, 46-52, 72-78, 115-121,
119-125, 260-266, 384-390, 459-465

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23, 232-243

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 451

CAACCATGCAAGGACAGGGCAGGAGAAGAGGAACCTGCAAAGACATATTTTGTTCCTCAAAATGG
CATCTTACCTTTATGGAGTACTCTTTGCTGTTGGCCTCTGTGCTCCAATCTACTGTGTGTCCC
CGGCCAATGCCCCCAGTGCATACCCCCGCCCTTCTCCACAAAGAGCACCCCTGCCTCACAGG
TGTATTCCCTCAACACCGACTTTGCCTTCCGCCTATACCGCAGGCTGGTTTTGGAGACCCCGA
GTCAGAACATCTTCTTCTCCCCTGTGAGTGTCTCCACTTCCCTGGCCATGCTCTCCCTTGGGG
CCCCTCAGTCACCAAGACCCAGATTCTCCAGGGCCTGGGCTTCAACCTCACACACACACCAG
AGTCTGCCATCCACCAGGGCTTCCAGCACCTGGTTCACTCACTGACTGTTCCAGCAAAGACC
TGACCTTGAAGATGGGAAGTGCCCTCTTCGTCAAGAAGGAGCTGCAGCTGCAGGCAAATTTCT
TGGGCAATGTCAAGAGGCTGTATGAAGCAGAAGTCTTTTCTACAGATTTCTCCAACCCCTCCA
TTGCCCAGGCGAGGATCAACAGCCATGTGAAAAAGAAGACCCAAGGGAAGGTTGTAGACATAA
TCCAAGGCCCTTGACCTTCTGACGGCCATGGTTCTGGTGAATCACATTTTCTTTAAAGCCAAGT
GGGAGAAGCCCTTTACCTTGAATATACAAGAAAGAACTTCCCATTCCTGGTGGGCGAGCAGG
TCACTGTGCAAGTCCCCATGATGCACCAGAAAGAGCAGTTCGCTTTTGGGGTGGATACAGAGC
TGAAGTGTCTTTGTGCTGCAGATGGATTACAAGGGAGATGCCGTGGCCTTCTTTGTCTCCCTA
GCAAGGGCAAGATGAGGCAACTGGAACAGGCCTTGTGAGCCAGAACACTGATAAAGTGGAGCC
ACTCACTCCAGAAAAGGTGGATAGAGGTGTTTCATCCCCAGATTTTCCATTTCTGCCTCCTACA
ATCTGGAAACCATCCTCCCGAAGATGGGCATCCAAAATGCCTTTGACAAAAATGCTGATTTTT
CTGGAATTGCAAAGAGAGACTCCCTGCAGGTTTCTAAAGCAACCCACAAGGCTGTGCTGGATG
TCAGTGAAGAGGGCACTGAGGCCACAGCAGCTACCACCACCAAGTTCATAGTCCGATCGAAGG
ATGGTCCCTCTTACTTCACTGTCTCCTTCAATAGGACCTTCTGATGATGATTACAAATAAAG
CCACAGACGGTATTCTCTTTCTAGGGAAAGTGGAAAATCCCCTAAATCCTAGGTGGGAAATG
GCCTGTAACTGATGGCACATTGCTAATGCACAAGAAATAACAAACCACATCCCTCTTTCTGT
TCTGAGGGTGCATTTGACCCCAAGTGGAGCTGGATTGCTGGCAGGGATGCCACTTCCAAGGCT
CAATCACCAAACCATCAACAGGGACCCCAAGTGCACAAGCCAACACCCATTAACCCCAAGTCAGTG
CCCTTTTCCACAAATTCTCCAGGTAAGTACTAGCTTCATGGGATGTTGCTGGGTACCATATTTT
CATTCCTTGGGGCTCCCAGGAATGGAATAACGCCAACCCAGGTTAGGCACCTCTATTGCAGAA
TTACAATAACACATTCAATAAACTAAAATATGAATTCAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

456/615

FIGURE 452

MASYLYGVLFVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLET
PSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVPSK
DLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQGKVVD
IIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAFGVDT
ELNCFVLQMDYKGDVAFFVLPSKKGMRQLEQALSARTLIKWSHSLQKRWIEVFIPRFSISAS
YNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTTKFIVRS
KDGPSYFTVSNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:
amino acids 1-20

20510"4352001

FIGURE 453

CTCCGGGTCCCCAGGGGCTGCGCCGGGCCGCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC
 CCCGCGGGGGCGATGACCGTGCGCTGACCCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCCGGGGCGACTCG
 GGGCGGACCGCGGGGCGGAGCTGCCGCCGTGAGTCCGGCCGAGCCACCTGAGCCCGAGCCGCGGGACACCGTC
 GCTCCTGCTCTCCGA**ATG**CTGCGCACCGCGATGGGCCTGAGGAGCTGGCTCGCCGCCCCATGGGGCGCGCTGCCG
 CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTCCTGCTGCAGCCGCCGCTCCGACCTGGGCGCTCAGC
 CCCCAGATCAGCCTGCCTCTGGGCTCTGAAGAGCGGCCATTCTCAGATTGAAGCTGAACACATCTCCAACCTAC
 ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC
 AACCTCAGCTTCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGACAGCGAGAGAAGAAACAGCAGTGC
 AGCTTCAAGGGCAAGGACCCACAGCGGACTGTCAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC
 CTGTTACCTGTGGCACAGCAGCCTTACGCCCCATGTGTACCTACATCAACATGGAGAACCTCACCCCTGGCAAGG
 GACGAGAAGGGGAATGTCTCTGGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC
 CTGGTGGTTGATGGCGAGCTTACACTGGAAAGTCAAGGGAATGACCCGGCCATCTCGCGGAGC
 CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC
 ATTCCTGAGAGCCTGGGCAGCTTGCAAGGCGATGATGACAAGATCTACTTTTTCTTCAGCGAGACTGGCCAGGAA
 TTTGAGTTCCTTTGAGAACACCATTTGTGTCCCGCATTGCCCGCATCTGCAAGGGCGATGAGGGTGAGAGCGGGTG
 CTACAGCAGCGCTGGACCTCCTTCCCTCAAGGCCAGCTGTGTGCTCACGGCCCGACGATGGCTTCCCTTCAAC
 GTGCTGCAGGATGTCTTACGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCCTTTTCTATGGGGTCTTCACT
 TCCCAGTGGCACAGGGGAACACAGAAGGCTCTGCCGTCTGTGTCTTCACAATGAAGGATGTGCAGAGAGTCTTC
 AGCGGCCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCCGTGCCACACCCCGG
 CCTGGAGCGTGATCACCAACAGTGCCCGGGAAGGAAGATCAACTCATCCCTGCAGCTCCCAGACCCGCTGCTG
 AACTTCTCAAGGACCACTTCTGTATGGACGGGAGGTCCGAAGCCGCATGCTGCTGCTGCAGCCCCAGGCTCGC
 TACCAGCGCTGGCTGTACACCGCGTCCCTGGCCTGCACACACCTACGATGTCTCTTCCCTGGGCACTGGTGAC
 GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCCGGGTGACATCATTTGAGGAGCTGCAGATCTTCTCATCGGGA
 CAGCCCGTGAGAACTCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCTCACACTCGGGCGTAGTCCAG
 GTGCCCATGGCCAACCTGCAGCCTGTACCGGAGCTGTGGGAGTGCCTCCTCGCCCGGACCCCTACTGTGCTTGG
 AGCGGCTCCAGCTGCAAGCACGTACGCTCTACAGCCTCAGCTGGCCACCAGGCCGTGGATCCAGGACATCGAG
 GGAGCCAGCGCCAAGGACCTTTGCAGCGCGTCTTCCGTTGTGTCCCCGTCTTTTGTACCAACAGGGGAGAGCCA
 TGTGAGCAAGTCCAGTTCCAGCCCAACACAGTGAACACTTTGGCCTGCCCCGTCTTCTCCACCTTGGCGACCCGA
 CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCCTCGGCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG
 GTGGGCACCCAACAGCTGGGGGAGTTCCAGTGTGGTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC
 TGCCAGAGGTGGTGGAGGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTCAATTATCAGCACA
 TCGCGTGTGAGTGACACAGCTGGTGGCAAGGCCAGCTGGGGTGACAGAGTCTTCTGGAAGGAGTTCTGGTG
 ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTATTCTTGCTCTACCGGCACCGGAACAGCATGAAA
 GTCTTCTGAAGCAGGGGGAATGTGCCAGCGTGACCCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC
 CCACTCAACGGCCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG
 GGGGCCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG
 TGCCCCCGGCCCGGGTCCGCTTGGCTCGGAGATCCGTGACTCTGTGGTG**TGA**GAGCTGACTTCCAGAGGACGC
 TGCCCTGGCTTACAGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCCCTCGCTCTGCTCTTCGTGGAAC
 ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC
 TACCACCCAGACACCCAAACAGCCGTGGCCCCAGAGGTCTGGCCAAATATGGGGGCTGCTTAGGTGGTGGA
 CAGTGCTCCTTATGTAACTGAGCCCTTTGTTTAAAAACAATTCCAAATGTGAAACTAGAATGAGAGGGAAGAG
 ATAGCATGGCATGCAGCACACAGGCTGCTCCAGTTTATGGCCTCCAGGGGTGCTGGGGATGCATCCAAAGTGG
 TTGTCTGAGACAGAGTTGGAAACCTCACCAACTGGCCTCTTACCTTCCACATTATCCCGCTGCCACCGGCTGC
 CCTGTCTCACTGCAGATTCAAGACCAGCTTGGGCTGCGTGCGTTCTGCTTGGCAGTCAGCCGAGGATGTAGTTG
 TTGCTGCCGTGCTCCACCACCTCAGGACCCAGAGGGCTAGGTGGCACTGCGGCCCTCACCAAGTCTGGGCTC
 GGACCCAACTCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGCGAGCTCAGGAGAGA
 TTTCTGTACAATGTACGCCTTTCCCTCAGAATTCAAGGAAGAGACTGTGCGCTGCTTCCCTCCGTTGTTGCGTGA
 GAACCCGTGTGCCCTTCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAAGTGCACC
 CTGGTCTCTCCCCAGTCCCCAGTTCAACCTCCATCCCTCACCTTCTTCCACTCTAAGGGATATCAACACTGCCC
 AGCACAGGGGCCCTGAATTTATGTGGTTTTTATAATTTTAAATAAGATGCACCTTATGTCAATTTTTTAATAAA
 GTCTGAAGAATTACTGTTTAAAAA

458/615

FIGURE 454

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962
><subunit 1 of 1, 837 aa, 1 stop
><MW: 92750, pI: 7.04, NX(S/T): 6
MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRFE
AEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEEKQOCSEFKGKD
PQRDCQNYIKILLPLSGSHLFTCGTAAFSMPCTYINMENFTLARDEKGNVLLEDGKGRCPPDP
NFKSTALVVDGELYTGTVSSSQNDPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPESLGS
LQGDDDKIYFFFSETGQEFEEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLCSRPD
GFPFNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSACVFTMKDVQRVFSGLYKEVN
RETQQWYTVTHPVPTPRPGACITNSARERKINSSLQLPDRVLNFLKDHFLMDGQVRSRMLLLQ
PQARYQRVAVHRVPGLHHTYDVLFGLTGDGRLHKAVSVGPRVHIIIEELQIFSSGQPVQNLLLD
THRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSLYQPQLATRPWIQ
DIEGASAKDLCSASSVVSPSFVPTGEKPCQVQFQPNVTNLTACPLLSNLATRLWLRNGAPVN
ASASCHVLPTGDLILLVGTQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQTDGGSVPV
STSRVSAPAGGKASWGADRSYWKFLVMCTLFVLAVLLPVLFLLYRHRNSMKVFLKQGECAV
HPKTCPPVLPPEPTRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKRPLSIQDSFVEVSP
VCPRPVRVLGSEIRDSVV

Transmembrane domains:

amino acids 23-46 (type II), 718-738

20510 9852500

FIGURE 455

TAAGATGAGGGCATCCCTCACGTTACACCCCCTGGTGGCATCTGCCAGCCCTGTTCTGGGGAC
AAGGCGGGCTTTTCGTGGGAGCCATGCTCAGCCTGCCAGGAAGCCAAGCCCTACAGTGCAGAGG
AAACAGAATTTCAACGGGAAGCTGGTTTGTCTTCATACCATTGGGATCTGCTGGTAAAGCTGTT
ATTTGGGTTTAGGGACTGATCCCTTGCACTTTACTTCTGGATCACCATGAATGGCCAAGATGG
TGGCAGAACACGCTGTGGACCCTGAGTTAGAGACAATGCAAATGTTGGATTGGGTGTAATTCT
TTTTGAATCCCAGATCCAGTCTGTACTTGAATATGAGCAGAAGATCTACAAGAATGCTGACAG
GGAACCGTGTTAAGACCCAGCACCCCTATTCCCAGGAGCTTCTGGCCTGACCATCTGCAGCCA
AAGCACTAACAGGGACAGATATGGGAATGTCCACCTTTGATCCGCATCCTGCACAATAGTGGT
CCCACCATGGCTGCCACTTTTTTATACTATTTGGAGAAAAGACCTTGTATAAATTTCGAGGCCC
GAGTGACTAACGTCTCTGTACACGGAATGGGTACTTGGTGGCATAGAGAAACACAATTAGC
CACTTTTTTCAGCTACACTTCTCACTCAGCTGCACCCCTACACTTCTCACTCAGGTGCACCCCT
TCTGCTGTCCTTTCCCAACGTACTGGGTCCCGAGCGTGGTGGGTATTTGCCACACTGGGTGC
CAGCTCAGCAGCCCCCACCTCTCTTTATTTCTCTCAAAGCTGGTCTTTCTGACTATCATTGT
GGTAGGGGGAGGACAGATGCTAAAGGTGGAAGCTGACCTGGAGAAAGAGACACACGGGGTGAC
TGTGGCAAAGGACAGCTGGAAAAGAACTCTATCACTTCTTCATTGGCAACCACAAGGCACCC
GAGGCCATGGCACTCCAGAGGCTGTGCGCAGAGCCAAGCCTCTCAACCTCTTCTGGCCCTGC
GTCCTGCAGCGAAGTCTCTGCTGTAAGACAGTAGACTCCTTCGATGAGGTGCTCAAAAATGCT
ACCCGGGGTGGTGGTGTCTGGCTTGCACTCTGGCCAGTTCAGAGAAAGTTGCAGAGATCAGGG
GCCAAGGATGTCATAGCCCCAGGTTGTCTCAGGGTCCCAATCCTAGGGCAGGGTGTGCATGG
AAGCAAGAACTATGGAACCTAGCTCCAGTCTGCAGGCTCTGAGCCCCTAGTTCCTCACTCCA
GCGGGGCTCCCTCACTGCACAGAACCCACCCCTTCTGTGTGGGCACTGCTGACCACACAGATG
ACCCAGACCCAAAGAGCCTGGCAGAAGCTCTGTGGTTGGAGCTGGGCTCCGTCTCCAGGTCTG
GTTTCAGGGGGATCAGGAAGGCTCTTTTCCACCTGTGGCTTCACTGGCCCTTTGAGATTTCCTA
TCTCACCGTTACTTCAGTTACCCTTGCAAGGGGGCCAGGGAGTCAAGAATATACCGTGTTCCTC
CAGGGTTTAAGCCGGCCATGCCTTCCCAGAGCATAACCAACTTGACAGGGGTGCCAGTTAC
CCCACAACTGAAGGAAGGAGATCCTTCCCCCGTCCCAGGAGTGCTCTCAACCAGCCTCAGA
AAGCTTGAGAAGATGGACCCTTTGCCACCAGGGTTAATTCTTGGTGGGGCAGCTCGGCTGTG
ATCAGGGCAACCAAACCTATAGGAAGCCTTCAGTGTGAGCTGGAATTAGACTGAACATGTGC
TTGGGCCTGCCTCTCCCTAGACGCAGTTGCGGGGCACTCCAGGGAATGAACCAGCTCAAGTGT
GTCCCTAACAGCAGCCTGGAGCTACCCCAATCCCTCACAGCCTGACCCTCCTCATTCCATCA
GATCTCGTGCCG

FIGURE 456

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA69555
><subunit 1 of 1, 148 aa, 1 stop
><MW: 16214, pI: 10.22, NX(S/T): 0
```

MGTWWHRETQLATFSATLLTQLHPTLLTQVHPLLLSFPQRTGSRWWVFATLGASSAAPH
LSLFSPKLVFLTIIIVGGGQMLKVEADLEKETHGVTVAKDSWKRNSITSSLATTRHPRPW
HSORLCAEPSLSTSSGPASCSEVSAVRQ

Important features of the protein:

Signal peptide:

Amino acids 1-28

Transmembrane domain:

Amino acids 64-78

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 103-107

N-myristoylation sites:

Amino acids 53-59; 94-100

CCGCGCGCGCCCTGGCACTCAATCCCCGCC**ATGT**GGGGGGCTCCTGCTCGCCCTGGCCGCCTTC
GCGCCGGCCGTCGGCCCCGGCTCTGGGGGCGCCAGGAACCTCGGTGCTGGGCCCTCGCGCAGCCC
GGGACCACCAAGGTCCCAGGCTCGACCCCGGCCCTGCATAGCAGCCCGGCACAGCCGCCGGCG
GAGACAGCTAACGGGACCTCAGAACAGCATGTCCGGATTTCGAGTCATCAAGAAGAAAAAGGTC
ATTATGAAGAAGCGGAAGAAGCTAACTCTAACTCGCCCCACCCCACTGGTGACTGCCGGGCC
CTTGTGACCCCACTCCAGCAGGGACCCTCGACCCCGCTGAGAAACAAGAAACAGGCTGTCTCT
CCTTTGGGTCTGGAGTCCCTGCGAGTTTCAGATAGCCGGCTTGAGGCATCCAGCAGCCAGTCC
TTTGGTCTTGACCACACCCGAGGACGGCTCAACATTTCATTTCAGGCCTGGAGGACGGCGATCTA
TATGATGGAGCCTGGTGTGCTGAGGAGCAGGACGCCGATCCATGGTTTCAGGTGGACGCTGGG
CACCCCAACCCGCTTCTCGGGTGTTATCACACAGGGCAGGAACTCTGTCTGGAGGTATGACTGG
GTCACATCATAACAAGGTCCAGTTCAGCAATGACAGTCGGACCTGGTGGGGAAGTAGGAACCAC
AGCAGTGGGATGGACGCAGTATTTCTCTGCCAATTTCAGACCCAGAAACTCCAGTGCTGAACCTC
CTGCCGGAGCCCCAGGTGGCCCGCTTCATTTCGCCTGCTGCCCCAGACCTGGCTCCAGGGAGGC
GCGCCTTGCCCTCCGGGCAGAGATCCTGGCCTGCCCAGTCTCAGACCCCAATGACCTATTCTCTT
GAGGCCCCCTGCGTCGGGATCCTCTGACCCTCTAGACTTTCAGCATCACAAATTACAAGGCCATG
AGGAAGCTGATGAAGCAGGTACAAGAGCAATGCCCAACATCACCCGCATCTACAGCATTGGG
AAGAGCTACCAGGGCCTGAAGCTGTATGTGATGGAAATGTTCGGACAAGCCTGGGGAGCATGAG
CTGGGGGAGCCTGAGGTGCGCTACGTGGCTGGCATGCATGGGAACGAGGCCCTGGGGCGGGAG
TTGCTTCTGCTCCTGATGCAGTTCCTGTGCCATGAGTTCTCTGCGAGGGAACCCACGGGTGACC
CGGCTGCTCTCTGAGATGCGCATTACACCTGCTGCCCTCCATGAACCTTGATGGCTATGAGATC
GCCTACCACCGGGGTTTCAGAGCTGGTGGGCTGGGGCCGAGGGCCGCTGGAACAACCAGAGCATC
GATCTTAACCATAAATTTTGCTGACCTCAACACACCACTGTGGGAAGCACAGGACGATGGGAAG
GTGCCCCACATCGTCCCCAACCATCACCTGCCATTGCCCACTTACTACACCTTGCCCAATGCC
ACCGTGGCTCCTGAAACGCGGGCAGTAATCAAGTGGATGAAGCGGATCCCCTTTGTGCTAAGT
GCCAACCTCCACGGGGGTGAGCTCGTGGTGTCTTACCCATTTCGACATGACTCGCACCCCGTGG
GCTGCCCGCGAGCTCACGCCCACACCAGATGATGCTGTGTTTCGCTGGCTCAGCACTGTCTAT
GCTGGCAGTAATCTGGCCATGCAGGACACCAGCCGCCGACCTGCCACAGCCAGGACTTCTCC
GTGCACGGCAACATCATCAACGGGGCTGACTGGCACACGGTCCCCGGGAGCATGAATGACTTC
AGCTACCTACACACCAACTGCTTTGAGGTCACTGTGGAGCTGTCTGTGACAAGTTCCTCAC
GAGAATGAATTGCCCCAGGAGTGGGAGAACAAACAAGACGCCCTCCTCACCTACCTGGAGCAG
GTGCGCATGGGCATTGCAGGAGTGGTGAGGGACAAGGACACGGAGCTTGGGATTGCTGACGCT
GTCATTGCCGTGGATGGGATTAACCATGACGTGACCACGGCGTGGGGCGGGGATTATTGGCGT
CTGCTGACCCAGGGGACTACATGGTGACTGCCAGTGCCGAGGGCTACCATTCAGTGACACGG
AACTGTTCGGGTACCTTTGAAGAGGGCCCCCTTCCCCTGCAATTTTCGTGCTCACCAAGACTCCC
AAACAGAGGCTGCGCGAGCTGCTGGCAGCTGGGGCCAAGGTGCCCCCGGACCTTCGCAGGCGC
CTGGAGCGGCTAAGGGGACAGAAGGAT**TGAT**ACCTGCGGTTTAAGAGCCCTAGGGCAGGCTGG
ACCTGTCAAGACGGGAAGGGGAAGAGTAGAGAGGGAGGGACAAAGTGAGGAAAAGGTGCTCAT
TAAAGCTACCGGGCACCTTAAA

462/615

FIGURE 458

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71162
><subunit 1 of 1, 734 aa, 1 stop
><MW: 81677, pI: 6.60, NX(S/T): 6
MWGLLLALAAFAPAVGPALGAPRNSVLGLAQPGTTKVPGSTPALHSSPAQPPAETANGTS
EQHVRIRVIKKKKVIMKKRKKLTLTRPTPLVTAGPLVTPTPAGTLDPAEKQETGCPPLGL
ESLRVSDSRLEASSSQSFGLGPHRGRLNIHSGLEDGDLYDGAWCAEEQDADPWFQVDAGH
PTRFSGVITQGRNSVWRYDWVTSYKVQFSNDSRTWWGSRNHSSGMDAVFPANSDPETPVL
NLLPEPQVARFIRLLPQTWLQGGAPCLRAEILACPVSDPNDLFLEAPASGSSDPLDFQHH
NYKAMRKLMKQVQECPNITRIYSIGKSYQGLKLYVMEMSDKPGHEHELGEPEVRYVAGMH
GNEALGRELLLLLLMQFLCHEFLRGNPRVTRLLSEMRIHLLPSMNPDGYEIAYHRGSELVG
WAEGRWNNQSIDLNHNFADLNTPLWEAQDDGKVPHIVPNHHLPLPTYTLPNATVAPETR
AVIKWMKRIPFVLSANLHGGELVVSYPFDMTRTPWAARELTPTPDDAVFRWLSTVYAGSN
LAMQDTSRRPCHSQDFSVHGNIINGADWHTVPGSMNDFSYLHTNCFEVTVELSCDKFPHE
NELPQEWENNKDALLTYLEQVRMGIAGVVRDKDTELGIADAVIAVDGINHDVTTAWGGDY
WRLLT PGDYMVTASAEGYHSVTRNCRVTFEEGPFPCNFVLTKTPKQRLRELLAAGAKVPP
DLRRRLERLRGQKD
```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

463/615

FIGURE 459

TAAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGAGAA
AGAACTGACTGAAACGTTTGAGATGAAGAAAGTTCTCCTCCTGATCACAGCCATCTTGGCAGT
GGCTGTTGGTTTCCCAGTCTCTCAAGACCAGGAACGAGAAAAAAGAAGTATCAGTGACAGCGA
TGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTGCCCCACTTCCACCAAT
TCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCCTATTCCAATACCTGAATCTGCCCC
TACAACCTCCCCTTCCTAGCGAAAAGTAAACAAGAAGGATAAGTCACGATAAACCTGGTCACCT
GAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCTGTTAATAAAAAGAAAAACAAAT
GTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCTTTAGTGATCTTCTTTAATAAACAT
GAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

2025-10-27 10:00:00

464/615

FIGURE 460

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290
><subunit 1 of 1, 85 aa, 1 stop
><MW: 9700, pI: 9.55, NX(S/T): 0
MKKVLLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPFRPLPPIPFPRFPW
FRRNFPIPIPIESAPTTPLPSEK

Important features of the protein:

Signal peptide:

amino acids 1-17

Homologous region to B3-hordein:

amino acids 47-85

Important features of the protein:

Signal peptide:

Amino acids 1-20

N-glycosylation sites:

Amino acids 57-61;210-214;220-224;318-322;428-432;472-476

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 80-84

N-myristoylation sites:

Amino acids 3-9;20-29;39-48;152-161;161-170;262-271;358-364;
538-544;560-566;637-643

Zinc carboxypeptidases, zinc-binding region 2 signature:

Amino acids 498-509

Zinc carboxypeptidases:

Amino acids 391-411

protein database

FIGURE 461

AGCAGGAGCAGGAGAGGGACAATGGGAAGCTGCCCCGTCCAGGTTTCATGTTCCCTCTTATTTCTC
CTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTGGT
GCTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCCACT
GAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCATAGC
ATGGTGCAAAAATTCCCAGGCGTGTCATTTGGGATCAGCACTGATTCTGAGGTTCTGACACAC
TACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACCTGAATTTA
GAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGATCAACAGC
CTCCACATGGTGACAGAGTACAACCTGTGACTGTGATTGGGTTATTCAACAGCGTAATTCAG
ATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATGCACAGATAC
CAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAGTGGTATGAAA
GAAAATGGGAAGGTGATATCATTTTTTCAAACATAAAGGAGTCTCAACTGCCAGCTTTGGCAATT
TACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCCGTAGAGCATGTG
CAAACTTTTGTGATGGATTCTTAAGTGGAAAATTGTTGAAAGAAAATCGTGAATCAGAAGGA
AAGACTCCAAAGGTGGAACCTTGACTTCTCCTTGGAACCTACATATGGCCAAGTATCTACTTTA
TGCAAAGTAAAAAGGCACAACTCAAATCTCAGAGACACTAAACAACAGGATCACTAGGCCTGC
CAACCACACACACACGCACGTGCACACACGCACGCACGCGTGCACACACACACGCGCACACAC
ACACACACACAGAGCTTCATTTCCCTGTCTTAAAATCTCGTTTTCTCTTCTTCTTTTAAA
TTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATTCTACTCTGTAAGCCCAT
CTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATGCCTCTATGAAAGAGAGGCA
TTCCTAGAGAAAGATTGTTCCAATTTGTCAATTAATATCAAGTTTGTATACTGCACATGACTT
ACACACAACATAGTTCCCTGCTCTTTTAAGGTTACCTAAGGGTTGAAACTCTACCTTCTTTTCA
AAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGGATGGTTTTAAACACCTTTGTGA
AATTGTCTTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTCCCTGAACTCAGCAGAAATAGACC
ATGTGAAAACCTCATGCTTGGTTAGCATCTCCAACCTCCCTATGTAAATCAACAACCTGCATAA
TAAATAAAAGGCAATCATGTTATA

466/615

FIGURE 462

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401
><subunit 1 of 1, 273 aa, 1 stop
><MW: 30480, pI: 4.60, NX(S/T): 1
MEAAPSRFMFLLFLLTCELAAEVAAEVEKSSDGPAAQEPTWLTDVPAAMEFIAATEVAVIGF
FQDLEIPAVPILHSMVQKFPGVSFGISTDSEVLTHYNITGNTICLFRLVDNEQLNLEDEDIES
IDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVIQIHLILLIMNKASPEYEENMHRYQKAALF
QGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLP TAEVSVEHVQNFC DGF
LSGKLLKENRESEGKTPKVEL
```

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

467/615

FIGURE 463

CTCGCTTCTTCCTTCTGGATGGGGGCCAGGGGGGCCAGGAGAGTATAAAGGCGATGTGGAGG
GTGCCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGGCC
ATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTATGGC
CCTGGAGGAGGCAAGTATTTTCAACCACTGAAGACTACGACCATGAAATCACAGGGCTGCGG
GTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAACTTGGAGACTCCTGGGACGTG
AAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATACATCACA
AAAGTCTTTGTGCGCTTCCAAGCTTTCTCCGGGGTATGGTCATGTACACCAGCAAGGACCGC
TATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAGAGGGGCAG
GTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGCTTTGAATGG
AATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTCAGCAAACCTCA
CCCGTGGGTGCGT**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGTGTGGTGGTGGC
TGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATAAATAAAGCTTCT
GCAGAAAA

2025-10-23 10:00:00

468/615

FIGURE 464

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSQVVKL
GDSWDVKLGALGGNTQEVTLQPGYITKVVFVAFQAFLRGMVMTSKDRYFYFGKLDGQISSAY
PSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

Signal peptide:
amino acids 1-22

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 465

CGGACGCGTGGGTCCGGCGGCCTGAGGCTGCACCGGGCACGGGTGCGCCGCAATCCAGCCTGGGCGGAGCCGGAG
 TTGCGAGCCGCTGCCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGGCCCGGAGAGCTCGGGGGACCCCGT
 TGCTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGACATATCCCTGGGC
 AGCCAGTCACCCCGCACTGGGTCTGGATGGACAACCCTGGCGCACCGTTCAGCCTGGAGGAGCCGGTCTCGAAGC
 CAGACATGGGGCTGGTGGCCCTGGAGGCTGAAGGCCAGGAGCTCCTGCTTGAGCTGGAGAAGAACCACAGGCTGC
 TGGCCCCAGGATACATAGAAACCCACTACGGCCCCAGATGGGCAGCCAGTGCTGCTGGCCCCCAACCACACGGATC
 ATTGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCTCTGACCTGCTCTGGGATGAGTG
 GCCTGATCACCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACTTCTCAA
 CCCACGAGATCTTTCGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGGCCACAGGGATCCTGGGAACAAAG
 CGGGCATGACCAGCCTTCTGGTGGTCCCCAGAGCAGGGGCAGGCGAGAAGCGCGCAGGACCCGGAAGTACCTGG
 AACTGTACATTGTGGCAGACCACACCCTGTTCTTGACTCGGCACCGAAACTTGAACCACACCAAACAGCGTCTCC
 TGGAAGTCGCCAACTACGTGGACCAGCTTCTCAGGACTCTGGACATTACAGGTGGCGCTGACCGGCTTGGAGGTGT
 GGACCGAGCGGGACCGCAGCCGCGTACGCAGGACGCCAACGCCACGCTCTGGGCCCTTCTGCACTGGCGCCGGG
 GGTGTGGGCGCAGCGGCCACGACTCCGCGCAGCTGCTCACGGGCCGCGCTTCCAGGGCGCCACAGTGGGCC
 TGGCGCCCGTGCAGGGCATGTGCGCGCCGAGAGCTCGGGAGGCGTGAGCACGGACCACTCGGAGCTCCCCATCG
 GCGCCGCAGCCACCATGGCCCATGAGATCGGCCACAGCCTCGGCCCTCAGCCACGACCCCGACGGCTGCTGCGTGG
 AGGCTGCGGGCCGAGTCCGGAGGCTGCGTTCATGGCTGCGGCCACCGGGCACCCGTTTCCGCGCGTGTTCAGCGCCT
 GCAGCCGCGGCCAGCTGCGCGCCTTCTTCCGCAAGGGGGGCGGCGCTTGCCCTCTCCAATGCCCGGACCCCGGAC
 TCCCGGTGCCCGCCGGCGCTCTGCGGGAACGGCTTCTGTGAAGCGGGCGAGGAGTGTGACTGCGGGCCCTGGCCAGG
 AGTGC CGGACCTCTGCTGCTTGTCTCACAAGTCTGCTGCGCCCGGGGGCCAGTGC CGCCACGGGGACTGCT
 GCGTGCCTGCTGCTGAAGCCGGCTGGAGCGCTGTGCGGCCAGGCCATGGGTGACTGTGACCTCCCTGAGTCTT
 GCACGGGCACCTCTCCCACTGTCCCCCAGACGTTTACCTACTGGACGGCTCACCTGTGCCAGGGGCAGTGGCT
 ACTGCTGGGATGGCGCATGTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCCCTGGCTCCACCCAGCTC
 CCGAGGCCTGTTTCCAGGTGGTGAACCTCTGCGGGAGATGCTCATGGAAGTGC CGGCCAGGACAGCGAGGGCCACT
 TCCGTGCCCTGTGCAGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCCAGCCTGCTCGCAC
 CGCACATGGTGCCAGTGGACTCTACCGTTACCTAGATGGCCAGGAAGTGACTTGTGCGGGAGCCTTGGCACTCC
 CCAGTGCCCGAGCTGGACCTGCTTGGCCCTGGGCCCTGGTAGAGCCAGGCACCCAGTGCTGGACCTAGAAATGGTGTGCC
 AGAGCAGGCGCTGCAGGAAGAATGCCCTTCCAGGAGCTTCAGCGCTGCCCTGACTGCCCTGCCACAGCCACGGGGTTT
 GCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCCCTTCTGTGACAAGCCAGGCTTTGGTGGCA
 GCATGGACAGTGGCCCTGTGCAGGCTGAAAACCATGACACCTTCTGCTGGCCATGCTCCTCAGCGTCTGCTGC
 CTCTGCTCCCAGGGGCCGGCTGGCCTGGTGTGCTACCGACTCCCAGGAGCCCATCTGCAGCGATGCAGCTGGG
 GCTGCAGAAGGGACCTTGCCTGCAGTGGCCCCAAAGATGGCCCCACAGGGACCACCCCCTGGGCGGCGTTCAAC
 CCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCTGGACCCCTGAGAACTCTCATGAGCCCGAGCAGCC
 ACCCTGAGAAGCCTCTGCCAGCAGTCTCGCCTGACCCCCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCT
 GGTGAGAGGTAGCTCCTAAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAACCTTGAAGT
 CAGGGGCAGAGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGGAAGTTTCTTCCCCGAGTGGAGCT
 TCGACCCACCCACTCCAGGAACCCAGAGCCACATTAGAAGTTTCTGAGGGCTGGAGAACTGCTTGGGCACACT
 CTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGGTACACAGCCCTGACCTCCCTCACCAGTGGAGGCTGG
 GTAGTGTGGCCATCCCCAAAGGGCTCTGCTCTGGGAGTCTGGTGTGCTCTCTACATGCAATTTCCACGGACCCCA
 GCTCTGTGGAGGGCATGACTGCTGGCCAGAAGCTAGTGGTCTTGGGGCCCTATGGTTCGACTGAGTCCACACTCC
 CCTGCAGCCTGGCTGGCCTCTGCAACAAACATAATTTTGGGGACCTTCTTCTCTGTTTCTTCCACCCCTGTCTT
 CTCCCCTAGGTGGTTCTTGAAGCCCCACCCCAATCCAGTGCTACACCTGAGGTTCTGGAGCTCAGAACTGAC
 AGCCTCTCCCCCATCTGTGTGTGTCCGGGGGACAGAGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAA
 AAGAAAGACATGTTGGCTATAGGCGTGGTGGCTCATGCCTATAATCCAGCACTTTGGGAAGCCGGGGTAGGAGG
 ATCACCAGAGGCCAGCAGGTCCACACCAGCCTGGGCAACAGCAAGACACCGCATCTACAGAAAAATTTTAAAAA
 TTAGCTGGGCGTGGTGGTGTGTACCTGTAGGCCTAGCTGCTCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTG
 AGTTCAACACTGCAGTGAGCTATGGTGGCACCCTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCTAA
 AATAAATTTTAAAGGACTTAAAAAAGAAAAAAGAAAAAAGAAAAA

470/615

FIGURE 466

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76788
><subunit 1 of 1, 813 aa, 1 stop
><MW: 87739, pI: 6.94, NX(S/T): 5

MGWRPRRARGTPLL LLLLLLLLLLWPVPGAGVLQGHIPGQPVTPHWVLDGQPWRTVSLEEPVSKPDMGLVALEAEGQ
ELLLELEKNHRL LAPGYIETHYGPDGQPVVLAPNHTDHCHYQGRVRGFPDSWVVLCTCSGMSGLITLSRNASYYL
RPWPPRGSKDFSTHEIFRMEQLLTWKGTGHRDPGNKAGMTSLPGGPQSRGRREARRTRKYLELYIVADHTLFLT
RHRNLNHTKQRLLEVANYVDQLLR TLDIQVALTGLEVWTERDRSRVTQDANATLWAF LQWRRGLWAQRPHDSAQL
LTGRAFQGATVGLAPVEGM CRAESSGGVSTDHSELP I GAAATMAHEIGHSLGLSHDPDGCCVEAAAESGGCVMAA
ATGHPFPFPRVFSACSRRLRAFFRKG GGA CLSNAPDPGLPVPPALCGNGFVEAGEECD CGPGQECRDLC CFAHNCS
LRPGAQCAHGDC CVRCLLK PAGALCRQAMGDCDLPEFCTGTSSHCPPDVYLLDGSPCARGSGYCWDGACPTLEQQ
CQQLWGPGSHPAPEACFQV VNSAGDAHNGCGQDSEGHFLPCAGRDALCGKLQCQGGKPSLLAPHMVPVDSTVHLD
GQEVTCRGALALPSAQLDLLGLGLVEPGTQCGPRMVCQSRRCRKNAFQELQRCLTACHSHGVCNSNHNCHCAPGW
APPFCDKPGFGGSGMDSGPVQAENHDTFLLAML LSVLLPLLPAGLAWCCYRLPGAHLQRCSWGCCRDPACSGPKD
GPHRDHPLGGVHPMELGPTATGQPWPLDPENSHEPSSHPEKPLPAVSPDPQADQVQMPRSC LW

Important features of the protein:

Signal peptide:

Amino acids 1-27

Transmembrane domain:

Amino acids 702-720

N-glycosylation sites:

Amino acids 109-113;145-149;231-235;276-280;448-452

Tyrosine kinase phosphorylation site:

Amino acids 236-244

N-myristoylation sites:

Amino acids 29-35;185-191;195-201;308-314;318-324;326-332;338-344;370-376;
400-406;402-408;454-460;504-510;510-516;517-523;580-586;
601-607;661-667;687-693;717-723;719-725

Amidation site:

Amino acids 200-204

Neutral zinc metallopeptidases, zinc-binding region signature:

Amino acids 342-352

20250101 09:56:01

[illegible]

472/615

FIGURE 468

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77623
><subunit 1 of 1, 97 aa, 1 stop
><MW: 10160, pI: 6.56, NX(S/T): 0
MQLGTGLLLAAVLSLQLAAAEAIWCHQCTGFGGCSHGSRCLRDSTHCVTTATRVLSNTEDLPL
VTKMCHIGCPDIPSLGLGPYVSIACCQTSLCNHD
```

Important features of the protein:

Signal peptide:

amino acids 1-20

N-myristoylation sites.

amino acids 6-11 and 33-38

Prokaryotic membrane lipoprotein lipid attachment sites.

amino acids 24-34 and 78-88

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 469

CATGGAGCCTCTTGCAGCTTACCCGCTAAATGTTCCGGGCCCAGAGCAAAGGTATTTGCAGT
TTTGCTGTCTATAGTTCTATGCACAGTAACGCTATTTCTTCTACAACCTAAAATTCCTCAAACC
TAAAATCAACAGCTTTTATGCCTTTGAAGTGAAGGATGCAAAAGGAAGAACTGTTTCTCTGGA
AAAGTATAAAGGCAAAGTTTCACTAGTTGTAAACGTGGCCAGTGACTGCCAACTCACAGACAG
AAATTACTTAGGGCTGAAGGAAGTGCACAAAGAGTTTGGACCATCCCACTTCAGCGTGTTGGC
TTTTCCCTGCAATCAGTTTGGAGAATCGGAGCCCCGCCCAAGCAAGGAAGTAGAATCTTTTGC
AAGAAAAAACTACGGAGTAACTTTCCCCATCTTCCACAAGATTAAGATTCTAGGATCTGAAGG
AGAACCTGCATTTAGATTTCTTGTTGATTCTTCAAAGAAGGAACCAAGGTGGAATTTTTGGAA
GTATCTTGTCAACCCTGAGGGTCAAGTTGTGAAGTTCTGGAGGCCAGAGGAGCCCATTTGAAGT
CATCAGGCCTGACATAGCAGCTCTGGTTAGACAAGTGATCATAAAAAAGAAAGAGGATCTATAG
AGAATGCCATTGCGTTTCTAATAGAACAGAGAAATGTCTCCATGAGGGTTTGGTCTCATTTTA
AACATTTTTTTTTTGGAGACAGTGTCTCACTCTGTCAACCAGGCTGGAGTGCAGTAGTGCGTT
CTCAGCTCATTGCAACCTCTGCCTTTTTTAAACATGCTATTAAATGTGGCAATGAAGGATTTTT
TTTTAATGTTATCTTGCTATTAAGTGGTAATGAATGTTCCCAGGATGAGGATGTTACCCAAAG
CAAAAATCAAGAGTAGCCAAAGAATCAACATGAAATATATTAATACTACTTCCTCTGACCATACT
AAAGAATTGAGAATACACAGTGACCAATGTGCCTCAATATCTTATTGTTCAACTTGACATTTT
CTAGGACTGTACTTGATGAAAATGCCAACACACTAGACCACTCTTTGGATTCAAGAGCACTGT
GTATGACTGAAATTTCTGGAATAACTGTAAATGGTTATGTTAATGGAATAAAACACAAATGTT
GAAAAATGTAAAATATATATACATAGATTCAAATCCTTATATATGTATGCTTGTTTTGTGTAC
AGGATTTTGTTTTTTCTTTTTAAGTACAGGTTCTAGTGTTTTACTATAACTGTCACTATGTA
TGTAAGTACATATATAAATAGTCATTTATAAATGACCGTATTATAACATTTGAAAAAGTCTT
CATCAAAAAAAAAAAAAA

474/615

FIGURE 470

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80136
><subunit 1 of 1, 209 aa, 1 stop
><MW: 23909, pI: 9.68, NX(S/T): 0
MEPLAAYPLKCSGPRAKVFAVLLSIVLCTVTLFLLQLKFLKPKINSFYAFEVKDAKGRTVSLE
KYKGKVSLLVNVASDCQLTDRNYLGLKELHKEFGPSHFSLAFPCNQFGESEPRPSKEVESFA
RKNYGVTTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNFWKYLVNPEGQVVKFWRPEEPIEV
IRPDIAALVRQVIIKKKEDL
```

Important features of the protein:

Signal peptide:

amino acids 1-31

Glutathione peroxidases signature 2.

amino acids 104-112

Glutathione peroxidases.

amino acids 57-82

20510"98525001

GCCCTAACCTTCCCAGGGCTCAGCTCTTTGGAGCTGCCATTTCCTCCGGCTGCGAGAAAGGACGCGCGCCCTGCC
TCGGGGCGAAGAAAAGAAGCAAAACCTTGTCTGGGAGGGTTTTCGTTCATCAACCTCCCTTCCCGCAAACCTAAACCTCCT
GCCGGGGCCATCCCTAGACAGAGGAAAGTTCCCTGCAGAGCCGACCAACCCCTAGTGGATCTGGGGCAGGCAGCGGG
GCTGGCTGTGGAATTAGATCTGTTTTGAACCCAGCTGGAGCGCATCGCTGGGGCTCGGAAGTCAACCGTCCCGCGGG
ACCGGGTTTGGCGCTGCCCGATGGAAACCGACAGTTTTCGAGCCCTCGGCTGCAAGTGGCCCTCTCCTCCCCCGGGTT
GTTTGTTCAGTGTCTGGGTGAGGGCTGCGAGTGTGGCAAGTTGCAAAGAGAGCCTCAGAGGTCCGAAGAGCGCTGCC
CTCCTACTCGCGTTTCGCTTCTTCTCTCTCGGTTCCCTACTGTGAAATCGCAGCGACATTTACAAAGGCCCTCCG
GGTCTTACCGAGACCGATCCGCAGCGTTTGGCCCCGGTCTGTGCCTATTGCATCGGGAGCCCCGAGCACCGGCGGAA
GGACTGGCGGGTGGGGTAGGGAGGTGGCGGGCGGGCGCATGGCGAGAGTTCCCGAAGGCCGACCTGGCCGCTGCAGG
AGTTATGTTACTTTGGCCACTTCTTTCACGGACCAGTTTCAGTTTCGCCGATGGGAAACCCCGGAGACCAAATCCTTGA
TTGGCAGTATGGAGTTTACTCAGGCCCTTCCCTCACACAGGAGGAGGTGGAAAGTTGATTCACACCGGTACAGCCA
CAGGTGGAAAAGAAACTTGGACTTTCTCAAGGCGGTAGACACGAACCGAGCAAGCGTCGGCCAAGACTCTCCTGA
GCCCAGAAGCTTCACAGACCTGCTGCTGGATGATGGGCAGGACAATAACACTCAGATCGAGGAGGATACAGACCA
CAATTACTATATATCTCGAATATATGGTCCATCTGATTCTGCCAGCCGGGATTTATGGGTGAACATAGACCAAAT
GGAAAAAGATAAAGTGAAGATTCATGGAAATATTGTCCAATACTCATCGGCCAAGCTGCAAGAGTGAATCTGTCCTT
CGATTTTTCCATTTTATGGCCACTTCTCTACGTGAAATCACTGTGGCAACCGGGGTTTATATACACTGGAGAAGT
CGTACATCGAATGCTAACAGCCACAGTACATAGCACCTTTAATGGCAAATTTTCGATCCCGAGTGTATCCAGAAA
TTCACTACTGTGAGATATTTTGATAATGGCACAGCACTTGTGGTCCAGTGGGACCATGTACATCTCCAGGATAATTA
TAACCTGGGAAGCTTCACATTCCAGGCAACCTTCTCATGGATGGACGAATCATCTTTGGATACAAAGAAATTC
TGTCTTGGTTCACACAGATAAGTTCAACCAATCATCCAGTGAAAGTCGGACTGTCCGATGCATTTGTCTGTGTCCA
CAGGATCCAAACAAATTTCCCAATGTGATTCGAAGAAGAACAAATTTATGAATACCACCCGAGTAGAGCTACAAATGTCAAA
AATTTACCAACATTTTCGGCTGTGGAGATGACCCATTACCCACATGCCCTCAGTTTAAACAGAGTGTGGCCCCCTGTGT
ATCTTCTCAGATTGGCTTCACTGCGAGTTGGTGTAGTAACTTCAAAGATGTTCCAGTGGATTGTATCGTCACTCG
GCAGGACTGGGTGGACAGTGGATGCCCTGAAGAGTCAAAAAGAGAAGATGTGTGAGAATACAGAACCAGTGGAAAC
TTCTTCTCGAACCACCACAACCGTAGGAGCGACAACCAACCCAGTTCAGGGTCTTAACCTACCACCCAGAGCAGT
GACTTCTCAGTTTCCCACAGCCTCCCTACAGAAGATGATACCAAGATAGCACTACATCTAAAAGATAATGGAGC
TTCTACAGATGACAGTGCAGCTGAGAAGAAGGGGGAACCTCCACGCTGGCCCTCATCTTGAATCCTCATCCT
GGTCTCATGTGAGCCCAAGCCATCTTGTGACAGTCTATATGTATCACCAACCAACATCAGCAGCCAGCATCTT
CTTTATGTAGAGACGCCCAAGCAGATGGCCTGCGATGAAGTTTAGAAGAGGCTCTGGACATCCTGCCCTATGCTGA
AGTTGAACCAGTTTGGAGAGAAAAGAAGGCTTTATTGTATCAGAGCAGTGCTAAATTTCTAGGACAGAACAACACC
AGTACTGGTTTTACAGGTGTTAAGACTAAAATTTTGCCTATACCTTTAAGACAAAACAAACAAACACACACAAAC
AAGCTCTAAGCTGCTGTAGCCTGAAGAAGACAAGATTTCTGGACAAGCTCAGCCACGGAACAAAGGGTAAACAA
AAAACATAAACTTATACAAGATACCAATTTACACTGAACATAGAATTTCCCTAGTGGAATGTATCTATAGTTCACT
CGGAACATCTCCCGTGGAATTTATCTGAAGTATGACAAGATTATAATGCTTTTGGCTTAGGTGCAGGGTTGCAAG
GGATCAGAAAAAATACTATAATAAAGCTTTAGTTTCATGAGG

FIGURE 472

MARFPKADLAAAGVMLLCHFFTDQFQFADGKPGDQILDWQYGVVTQAFPHTEEEVEVDSHAYSH
RWKRNLDFLKAVDTNRA SVGQDSPEPRSFTDLLDDGQDNNTQIEEDTDHNYISRIYGPSDS
ASRDLWVNIDQMEKDKVKIHGILSNTHRQAARVNLSFDFPFYGHFLREITVATGGFIYTGEVV
HRMLTATQYIAPLMANFDPVS SRNSTVRYFDNGTALVVQWDHVHLQDNYNLGSFTTFQATLLMD
GRIIFGYKEIPVLVTQISS TNHPVKVGLSDAFVVVHRIQQIPNVRRRTIY EYHRVELQMSKIT
NISAVEMTPLPTCLQFNRCGPCVSSQIGFNCSWCSKLQRCSSGFD RHRQDWVDSGCPEESKEK
MCENTEPVETSSRTTTTVGATTTQFRVLTTTTRAVTSQFPTSLPTEDDTKIALHLKDNASTD
DSAAEKKGGTLHAGLIIGILILVLIVATAILVTVYMYHHPTSAASIFFIERRPSRWPAMKFRR
GSGHPAYAEVEPVGEKEGFIVSEQC

Important features of the protein:**Transmembrane domain:**

amino acids 454-478

N-glycosylation sites.

amino acids 103-107, 160-164, 213-217, 221-225, 316-320, 345-349

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 297-301, 492-496, 503-507

N-myristoylation sites.

amino acids 42-48, 100-106, 147-153, 279-285, 397-403, 450-456,
455-461

FIGURE 473

CGCGGAGCCCTGCGCTGGGAGGTGCACGGTGTGCACGCTGGACTGGACCCCCATGCAACCCCG
CGCCCTGCGCCTTAACCAGGACTGCTCCGCGCGCCCTGAGCCTCGGGCTCCGGCCCGGACCT
GCAGCCTCCAGGTGGCTGGGAAGAACTCTCCAACAATAAATACATTTGATAAGAAAAGATGGC
TTTAAAAGTGCTACTAGAACAAGAGAAAACGTTTTTCACTCTTTTAGTATTACTAGGCTATTT
GTCATGTAAAGTGACTTGTGAATCAGGAGACTGTAGACAGCAAGAATTCAGGGATCGGTCTGG
AAACTGTGTTCCCTGCAACCAGTGTGGGCCAGGCATGGAGTTGTCTAAGGAATGTGGCTTCGGC
TATGGGGAGGATGCACAGTGTGTGACGTGCCGGCTGCACAGGTTCAAGGAGGACTGGGGCTTC
CAGAAATGCAAGCCCTGTCTGGACTGCGCAGTGGTGAACCGCTTTCAGAAGGCAAATTTGTTCA
GCCACCAGTGATGCCATCTGCGGGGACTGCTTGCCAGGATTTTATAGGAAGACGAAACTTTGTC
GGCTTTCAAGACATGGAGTGTGTGCCTTGTGGAGACCCTCCTCCTCCTTACGAACCGCACTGT
GCCAGCAAGGTCAACCTCGTGAAGATCGCGTCCACGGCCTCCAGCCCACGGGACACGGCGCTG
GCTGCCGTTATCTGCAGCGCTCTGGCCACCGTCCTGCTGGCCCTGCTCATCCTCTGTGTCTATC
TATTGTAAGAGACAGTTTTATGGAGAAGAAACCCAGCTGGTCTCTGCGGTTCGACGACATTTCAG
TACAACGGCTCTGAGCTGTCTGTTTTGACAGACCTCAGCTCCACGAATATGCCCACAGAGCC
TGCTGCCAGTGCCGCCGTGACTCAGTGCAGACCTGCGGGCCGGTGCCTTGCTCCCATCCATG
TGCTGTGAGGAGGCCTGCAGCCCCAACCCGCGCACTCTTGTTGTGGGTGCATTCTGCAGCC
AGTCTTCAGGCAAGAAACGCAGGCCCAGCCGGGGAGATGGTGCCGACTTCTTCGGATCCCTC
ACGCAGTCCATCTGTGGCGAGTTTTTCAGATGCCTGGCCTCTGATGCAGAATCCCATGGGTGGT
GACAACATCTCTTTTTGTGACTCTTATCCTGAACTCACTGGAGAAGACATTCACTCTCTCAAT
CCAGAACTTGAAAGCTCAACGTCTTTGGATTCAAATAGCAGTCAAGATTTGGTTGGTGGGGCT
GTTCCAGTCCAGTCTCATTCTGAAAACCTTTACAGCAGCTACTGATTTATCTAGATATAACAAC
ACACTGGTAGAATCAGCATCAACTCAGGATGCACTAACTATGAGAAGCCAGCTAGATCAGGAG
AGTGGCGCTGTCTACCCAGCCACTCAGACGTCCCTCCAGGAAGCTTAAAGAACCTGCTTC
TTTCTGCAGTAGAAGCGTGTGCTGGAACCCAAAGAGTACTCCTTTGTTAGGCTTATGGACTGA
GCAGTCTGGACCTTGATGGCTTCTGGGGCAAAAATAAATCTGAACCAAACCTGACGGCATTTC
AAGCCTTTCAGCCAGTTGCTTCTGAGCCAGACCAGCTGTAAGCTGAAACCTCAATGAATAACA
AGAAAAGACTCCAGGCCGACTCATGATACTCTGCATCTTTCCTACATGAGAAGCTTCTCTGCCAC
AAAAGTGACTTCAAAGACTGATGGGTGAGCTGGCAGCCTATGAGATTGTGGACATATAACAA
GAAACAGAAATGCCCTCATGCTTATTTTCATGGTGATTGTGGTTTTTACAAGACTGAAGACCCA
GAGTATACTTTTTCTTTCCAGAAATAATTTATACCGCCTATGAAATATCAGATAAATTACCT
TAGCTTTTTATGTAGAATGGGTTCAAAGTGAGTGTCTTCTATTTGAGAAGGACACTTTTTTCATC
ATCTAAACTGATTTCGCATAGGTGGTTAGAATGGCCCTCATATTGCCTGCCTAAATCTTGGGTT
TATTAGATGAAGTTTACTGAATCAGAGGAATCAGACAGAGGAGGATAGCTCTTTCAGAAATCC
ACACTTCTGACCTCAGCCTCGGTCTCATGAACACCCGCTGATCTCAGGAGAACACCTGGGCTA
GGGAATGTGGTTCGAGAAAGGGCAGCCCATGCCCAGAATTAACACATATTGTAGAGACTTGTA
TGCAAAGGTTGGCATATTTATATGAAAATTAGTTGCTATAGAAACATTTGTTGCATCTGTCCC
TCTGCCTGAGCTTAGAAGGTTATAGAAAAAGGGTATTTATAAACATAAATGACCTTTTACTTG
CATTGTATCTTATACTAAAGGCTTTAGAAATTACAACATATCAGGTTCCCCTACTACTGAAGT
AGCCTTCCGTGAGAACACACCACATGTTAGGACTAGAAGAAAATGCACAATTTGTAGGGGTTT
GGATGAAGCAGCTGTAACCTGCCCTAGTGTAGTTTGACCAGGACATTGTCTGCTCCTTCCAAT
TGTGTAAGATTAGTTAGCACATCATCTCCTACTTTAGCCATCCGGTGTGGATTTAAGAGGAC
GGTGCTTCTTTCTATTAAAGTGCTCCATCCCCTACCATCTACACATTAGCATTGTCTCTAGAG
CTAAGACAGAAATTAACCCGTTTCAGTCAAAAGCAGGGAATGGTTCACTTACTCTTAATCTT
TATGCCCTGGAGAAGACCTACTTGAACAGGGCATATTTTTTTAGACTTCTGAACATCAGTATGT
TCGAGGGTACTATGATATTTTTGGTTTGAATTGCCCTGCCCAAGTCACTGTCTTTTAACTTTT
AAACTGAATATTAAATGTATCTGTCTTTCCT

478/615

FIGURE 474

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84210
><subunit 1 of 1, 417 aa, 1 stop
><MW: 45305, pI: 5.12, NX(S/T): 6
MALKVLLLEQEKTFFTLLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK
ECGFGYGEDAQCVTCLHRFKEDWGFQKCKPCLDCAVVNRFQKANC SATSDAICGDCLPG
FYRKTCLVGFQDMECVPCGDP PPPYPHCASKVNLVKIASTASSPRDTALAAVIC SALAT
VLLALLILCVIYCKRQFMEEKP SWSLRSQDIQYNGSELSCFDRPQLHEYAHRACCQCRRD
SVQTCGPVRL LPSMCCEEACSPNPATLGCGVHSAASLQARNAGPAGEMVPTFFGSLTQSI
CGEFS DAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV
PVQSHSENFTAATDL SRYNNTLVESASTQDALTMRSQLDQESGAVIHPATQTSLQEA

Important features of the protein:

Signal peptide:

Amino acids 1-25

Transmembrane domain:

Amino acids 169-192

N-glycosylation sites:

Amino acids 105-109;214-218;319-323;350-354;368-372;379-383

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 200-204;238-242

Tyrosine kinase phosphorylation site:

Amino acids 207-214

N-myristoylation sites:

Amino acids 55-61;215-221;270-276

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 259-270

TNFR/NGFR family cysteine-rich region proteins:

Amino acids 89-96

478/615 " 98525001

FIGURE 475

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATGG**CAATG
GGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCCAA
GCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCCTTTGACCTCAGGGCCATGAAG
CTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGCTATCCTGTCTCTGGACACCGGGTAT
CGCGCGCCGGTGACCCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGACGCAA
TCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACTGACAAATGC
AACGCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCCGCCGACG
CTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTATCGGCAGG
TCCCGACGAGTCCAGTGTCAACAGGACCAGACCGCCTGCTTCCAGGGCAGTGGCAGAATGACA
GTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTGCACCACCGAG
GGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGGGGTACCTCTGC
AACAGGAAATCCATGACCCAGCCCTTCACCAGTGCTTCAGCCACCACCCCTCCCCGAGACTA
CAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGCA**TAG**ACCGCCCCT
CCAGGATGCTGGGGACAGGGCTCACACACCTCATTCTTGCTGCTTCAGCCCCTATCACATAGC
TCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

480/615

FIGURE 476

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576
><subunit 1 of 1, 251 aa, 1 stop
><MW: 26935, pI: 7.42, NX(S/T): 2
MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSLD
TGYRAPVTILVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQAPD
PPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQGSGRMTVGNFSVPVYIRTCHRPSC
TTEGTTSPWTAIDLQGSCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVGLSA
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

10055555.0150

CCACACGCGTCCGGGACAGATGAACCTTAAAAGAGAAGCTTTAGCTGCCAAAGATTGGGAAAAGGAAAAGGACAAAA
AGACCCCTGGGCTACACGGCGTAGGTGCAGGGTTTCCCTACTGCTGTTCTTTTATGCTGGGAGCTGTGGCTGTAAC
CAACTAGGAAATAACGTATGCAGCAGCTATGCGCTGTCTAGAGAGTTGTGCTTCCCAAGACAAAGGCAAGTCCTGTT
TCTTTTTTCTTTTTTGGGGAGTGTCTTGGCAGGTTCTGGGTTTGGACGTTATTCCGGTGACTGAGGAAAAAGAGAA
AGGATCCTTTGTGGTCAATCTGGCAAAAGGATCTGGGACTAGCAGAGGGGGAGCTGGCTGCAAGGGGAAACCAAGGTT
GGTTTCCGATGATAACAAACAATACTGCTCTGGATTACATACCGGGAATTTTGCTCAAAATGAGAAATCGGA
CCGAGAGAAGCTGTGTGGCCCTAAAGAGCCCTGTATGCTGTATTTCCAAATTTTAATGGATGATCCCTTTTCAGAT
TTACCGGGCTGAGCTGAGAGTCAGGGATATAAATGATCACGCGCCAGTATTTTTCAGGACAAAGAAACAGTCTTAA
AATATCAGAAAAATACAGCTGAAGGGACAGCATTTAGACTAGAAAGAGCACAGGATCCAGATGGAGGACTTAACGG
TATCCAAAACCTACACGATCAGCCCCAACTCTTTTTTCCATATTAACATGAGTGGCGGTGATGAAGGCATGATATA
TCCAGAGCTAGTGTGGACAAAGCACTGGATCGGGAGGAGCAGGGAGAGCTCAGCTTAACCTTCACAGCGCTGGA
TGGTGGGTCTCCATCCAGGTCTGGGACCTCTACTGTACGCATCGTTGTCTTGAGACGTCAATGACAATGCCCCACA
GTTTGCCAGGCTCTGTATGAGACCCAGGCTCCGAAAAACAGCCCCATTGGGTTCTTTATTGTTAAGGTATGGGC
AGAAGATGTAGACTCTGGAGTCAACGCGGAAGTATCCTATTCATTTTTTGATGCCTCAGAAAAATATTCGAACGAC
CTTTCAAATCAATCCTTTTTCTGGGAAATCTTTCTCAGAGAATTGCTTGATTATGAGTTAGTAAATCTTTACAA
AATAAATATACAGGCAATGGACGGTGGAGGCCCTTCTGCAAGATGTAGGGTTTATGTTGAAGTATTGACACCCAA
TGACAAATCCCCCTGAACATGATCGTATCATATTTTCCAACCTGTGTGCTGAGAAATTCCTCTGAGACGCCGCTGGC
TGTTTTTAAGATTAATGACAGACTCTGGAGAAAAAGAAAGATGGTTTGCCTACATCAAGAGAATCTGCCATT
CCTACTAAAAACCTTCTGTGGAGAATTTTTACATCCTAATTAAGAAGCGCGCTGGACAGAGAGATCAGAGCCGA
GTACAACATCACTATCACCGTCACTGACTTTGGGGACACCCAGGCTGAAAACCGAGCACAAACATAACGGTCTTGGT
CTCCGACGTCAATGACAACGCCCCCGCCTTACCCAAACCTCCTACACCTGTTGTCGCGGAGAACAAACAGCCC
CGCCCTGCACATCGGCAGCGTCAGCGCCACAGACAGAGACTCGGGCACCAACGCCCAGGTCACTACTCGCTGCT
GCCGCCCAAGACCCGCACCTGCCCTCGCCTCCTCGCTCCTCATCAACGCGGACAAACGGCCACCTGTTTCGCCCT
CAGGTCTGCTGAGCTACGAGGCCCTGTCAGGCTTTCGAGTTCCGCGTGGGCGCCACAGACCGCGGCTCCCCCGCGCT
GAGCAGACAGGCGCTGGTGCCTGCTGGTCTGAGACGCCAACGACAACTCGCCCTCGTGTGATACCCGCTGCA
GAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCTCGGGCGGCGGAGCCGGGCTACCTGGTGACCAAGGTGGTGGC
GGTGGACGGCGACTCGGGCCAGAACGCCCTGGCTGTCTGATACAGCTGCTCAAGGCCACGGAGCCCGGGCTGTTTCGG
TGTGTGGGCGCACAAATGGGGAGGTGCGCACCGCCAGGCTGCTGAGCGAGCGCGACGCAGCCAAAGCACAGGCTCGT
GGTGTCTTGTCAAAGACAATGGCGAGCCTCCTCGCTCGGCCACCGCCACGCTGCATCTTGCTCCTGGTGGACGGCTT
CTCCCAGCCCTACCTGCTCTCCCGAGGGCGGCCCGGCCACAGGCCAGGCCGAGGCCGAGCTGCTCACCCTCTA
CCTGGTGTGGCTTGGCTCGGTGCTCTCGCTCTTCTCTCTCGTGCTCCTGTCTGTTGCTGGCGGTGCGGCTGTG
CAGGAGGAGCAGGGCGGCTCGGTGGGTGCTGCTCGGTGCCCGAGGGTCTTTTTTCAGGGCATCTGGTGGACGT
GAGGGGCGCTGAGACCTGTCCAGAGCTACCAGTATGAGGTGTGCTGACGGGAGGCCCCGGGACCAGTGAGTT
CAAGTTCTTGAAACAGTTATTTTCGATATTTCAGGCACAGGGCCCTGGGAGGAAGGGTGAAGAAAAATTCACCTT
CCGAAATAGCTTTGGATTTAATATTTCAGTAAAGTCTGTTTTTATGTTTCATATACTTTTTGGTGTGTTACATAGCCA
TGTTTTCTATTAGTTTACTTTTTAAATCTCAAAATTTAAGTTATTATGCAACTTCAAGCATATTTTTCAAGTAGTATA
CCCCGTGGTTTTTCAACTGTTTTCATATTTTTTTCGATTAATAACAACTGGGTTTAATTTAATGAGTATTTTTT
CTAAATGATAGTGTTAAGGTTTTAATTTCTTCCAACTGCCCAAGGAATTAATTTACTATTATATCTCATTAACAGAA
ATCTGAGGTTTTTGATTCAATTCAGAGCTTGCATCTCATGATTCTAATCACTTCTGTCTATAGTGTACTTTGCTCTA
TTTAAGAAGGCATATCTACATTTCCAAACTCATTCTAACATTCTATATATTCTGTGTTTGAAAACCATGTCATTTA
TTTCTACATCATGTATTTAAAAAGAAATATTTCTCTACTACTATGCTCATGACAAAATGAAAACAAAGCATATTGT
GAGCAATACTTGAACATCAATAATACCCCTTAGTTTATATACTATTATTTTATCTTTAAGCATCTACTTTTACTT
GGCCAAATATTTTTCTTATGTTTAACTTTTTGCTGATGTATAAAACAGACATGCTCTATAATGAAATAAAATTATAA
TGCCCTTGAAAAATGAATAAAAAATAAAACATTTTGAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

482/615

FIGURE 478

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA87976
><subunit 1 of 1, 800 aa, 1 stop
><MW: 87621, pI: 4.77, NX(S/T): 7
MAVRELCFPRQRQVLFLFLFWGVSLAGSGFGRYSVTEETEKGSFVVNLAKDLGLAEGELAARG
TRVVSDDNKQYLLLDSTHTGNLLTNEKLDREKLCGPKEPCMLYFQILMDDPFQIYRAELRVLDI
NDHAPVFQDKETVLKISENTAEGTAFRLERAQDPDGLNGIQNYTISPNSFFHINISGGDEGM
IYPELVLDKALDREEQGELSLLTALDGGSPSRSGTSTVRIVVLDVNDNAPQFAQALYETQAP
ENSPIGFLIVKVWAEDVDSGVNAEVSYSFFDASENIRTTFQINPFSGEIFLRELLDYELVNSY
KINIQAAMDGGGLSARCRVLVEVLDTNDNPPELIVSSFSNSVAENSPETPLAVFKINDRDSGEN
GKMVCYIQENLPFLKPSVENFYILITEGALDREIRAENITITVTDLGTPRLKTEHNITVLV
SDVNDNAPAFQTQTSYTLFVRENNSPALHIGSVSATDRDSGTNAQVTYSLLPPQDPHLPLASLV
SINADNGHLFALRSLDYELQAFEFVRGATDRGSPALSREALVRVLVLDANDNSPFVLYPLQN
GSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATEPGLFGVWAHNGEVRTARLL
SERDAAKHRLVVLVKDNGEPPRSATATLHLLLVLDGFSQPYLPLPEAAPAQAEADLLTVYLV
VALASVSSLFLLSVLLFVAVRLCRRSRAASVGRCSVPEGPFPGHLVDVRGAETLSQSYQYEV
LTGGPGTSEFKFLKPVISDIQAQGPGRKGEENSTFRNSFGFNIQ

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 687-711

N-glycosylation sites.

amino acids 169-173, 181-185, 418-422, 436-440, 567-571, 788-792

Glycosaminoglycan attachment site.

amino acids 28-32

Tyrosine kinase phosphorylation sites.

amino acids 394-402, 578-585

N-myristoylation sites.

amino acids 22-28, 27-33, 53-59, 82-88, 162-168, 184-190,
217-223, 324-330, 325-331, 471-477, 568-574, 759-765

Amidation site.

amino acids 781-785

Aminoacyl-transfer RNA synthetases class-II signature 1.

amino acids 117-138

Cadherins extracellular repeated domain signature.

amino acids 121-132, 230-241, 335-346, 439-450, 549-560

FIGURE 479

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCCTGGGAATTTAAGGGACCCACACTACCTTCCC
GAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTCA**ATG**GGACCTGTGCGGTTGG
GAATATTGCTTTTTCTTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGAAGGAGGAGG
ACGATGACACAGAACGCTTGCCCGAGCAAATGCGAAGTGTGTAAGCTGCTGAGCACAGAGCTAC
AGGCGGAAGTGAAGTGCACCGGTGATCTCGAGAGGTGCTGGAGCTGGGGCAGGTGCTGGATA
CAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAAGGCTGGAAGAGGCCTTAG
AGAATTTATGTGAGCGGATCCTGGACTATAGTGTTTACGCTGAGCGCAAGGGCTCACTGAGAT
ATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAGGCCTAGTGCAGAAGGGGGTGAAGG
TGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGCCCAGCGTGGAGGTCACATACCTCAAGA
AGCAGTGTGAGACCATGTTGGAGGAGTTTGAAGACATTGTGGGAGACTGGTACTTCCACCATC
AGGAGCAGCCCCTACAAAATTTTCTCTGTGAAGGTCATGTGCTCCCAGCTGCTGAAACTGCAT
GTCTACAGGAACTTGGACTGGAAAGGAGATCACAGATGGGGGAAGAGAAAACAGAAGGGGAGG
AAGAGCAGGAGGAGGAGGAGGAAGAGGAGGAAGAGGAAGGGGGAGACAAGATGACCAAGACAG
GAAGCCACCCCAAACCTTGACCGAGAAGATCTTT**TGA**CCCTTGCCTTTGAGCCCCCAGGAGGGGA
AGGGATCATGGAGAGCCCTCTAAAGCCTGCACTCTCCCTGCTCCACAGCTTTCAGGGTGTGTT
TATGAGTGAATCCACCAAGCTTGTAGCTGTTCTCTCCCATCTAACCTCAGGCAAGATCCTGG
TGAAACAGCATGACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGGTCCTGCTCCTAGAGATG
AACTCTATCCAGCCCCCTTAATTGGCAGGTGTATGTGCTGACAGTACTGAAAGCTTTCCTCTTT
AACTGATCCCACCCCAACCAAAAGTCAGCAGTGGCACTGGAGCTGTGGGCTTTGGGGAAAGTC
ACTTAGCTCCTTAAGGTCTGTTTTTTAGACCTTCCAAGGAAGAGGCCAGAACGGACATTCTCT
GCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCAGCAAACCGTGAAGGAGAATGG
GACACTGGGTGATGGCCTGGAGTTGCTGATAATTTAGGTGGGATAGATACTTGGTCTACTTAA
GCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGTGCTCAACTTTCTATATCGCTATT
AAACTTTTTTCTTTTTTTCTA

484/615

FIGURE 480

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92256
><subunit 1 of 1, 248 aa, 1 stop
><MW: 28310, pI: 4.63, NX(S/T): 0
MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSKCEVCKLLSTELQAE LSRTGRSREVLE
LGQVLDTGKRKRHVPYSVSETRLEEALENLCERILDYSVHAERKGS LRYAKGQSQTMATLKGL
VQKGVKVDLGIPLELWDEPSVEV TYLKKQCETMLEEFEDIVGDWYFHHQE QPLQNFLCEGHVL
PAAETACLQETWTGKEITDGE EKTEGEEEEEQEEEEEEEEEGGDKMTKTGSHPKLDREDL
```

Important features of the protein:

Signal peptide:

amino acids 1-21

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 106-110

N-myristoylation site.

amino acids 115-121

Amidation site.

amino acids 70-74

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 481

GGCGTGTGCAAGGCGGGGTCCGGCCCCGCGCAGGTCTGGGTAAAGCGCGTCTAGGGCGCTGCGCGG
CGCAGCGAAAA**ATG**GCGGCTTCCAGGTGGGCGCGCAAGGCCGTGGTCCTGCTTTGTGCCTCTGA
CCTGCTGCTGCTGCTGCTACTGCTACCACCGCCTGGGTCTGCGCGGCCGAAGGCTCGCCCCG
GACGCCCCGACGAGTCTACCCACCTCCCCGGAAGAAGAAGAAGGATATTGCGGATTACAATGA
TGCAGACATGGCGCGTCTTCTGGAGCAATGGGAGAAAGATGATGACATTGAAGAAGGAGATCT
TCCAGAGCACAAGAGACCTTCAGCACCTGTGACTTCTCAAAGATAGACCCAAGCAAGCCTGA
AAGCATATTGAAAATGACGAAAAAAGGGAAGACTCTCATGATGTTTGTCACTGTATCAGGAAG
CCCTACTGAGAAGGAGACAGAGGAAATTACGAGCCTCTGGCAGGGCAGCCTTTTCAATGCCAA
CTATGACGTCCAGAGGTTTATTGTGGGATCAGACCGTGCTATCTTCATGCTTCGCGATGGGAG
CTACGCCTGGGAGATCAAGGACTTTTTGGTTCGGTCAAGACAGGTGTGCTGATGTAACCTCTGGA
GGGCCAGGTGTACCCCGGCAAAGGAGGAGGAAGCAAAGAGAAAAATAAAACAAAGCAAGACAA
GGGCAAAAAAAGAAGGAAGGAGATCTGAAATCTCGGTCTTCCAAGGAAGAAAATCGAGCTGG
GAATAAAAGAGAAGACCTG**TGAT**GGGGCAGCAGTGACGCGCTGTGGGGGGACAGGTGGACGTG
GAGAGCTCTTTGCCAGCTCCTGGGGTGGGAGTGGTCTCAGGCAACTGCACACCGGATGACAT
TCTAGTGTCTTCTAGAAAGGGTCTGCCACATGACCAGTTTGTGGTCAAAGAATTACTGCTTAA
TAGGCTTCAAGTAAGAAGACAGATGTTTTCTAATTAATACTGGACACTGACAAATTCATGTTT
ACTATAAAATCTCCTTACATGGAAATGTGACTGTGTTGCTTTTTCCCATTTACACTTGGTGAG
TCATCAACTCTACTGAGATTCCACTCCCCTCCAAGCACCTGCTGTGATTGGGTGGCCTGCTCT
GATCAGATAGCAAATTCTGATCAGAGAAGACTTTAAAACTCTTGACTTAATTGAGTAAACTCT
TCATGCCATATACATCATTTTCATTATGTTAAAGGTAAAATATGCTTTGTGAACTCAGATGTC
TGTAGCCAGGAAGCCAGGGTGTGTAAATCCAAAATCTATGCAGGAAATGCGGAGAATAGAAAA
TATGTCACTTGAAATCCTAAGTAGTTTTGAATTTCTTTGACTTGAATCTTACTCATCAGTAAG
AGAACTCTTGGTGTCTGTGAGGTTTTATGTGGTCTGTAAAGTTAGGGGTTCTGTTTTGTTTTCC
TTATTTAGGAAAGAGTACTGCTGGTGTGAGGGGTTATATGTTCCATTTAATGTGACAGTTTTT
AAAGGATTTAAGTAGGGAATCAGAGTCCTTTGCAGAGTGTGACAGACGACTCAATAACCTCAT
TTGTTTCTAAACATTTTTCTTTGATAAAGTGCCTAAATCTGTGCTTTCGTATAGAGTAACATG
ATGTGCTACTGTTGATGTCTGATTTTGCCGTTTCATGTTAGAGCCTACTGTGAATAAGAGTTAG
AACATTTATATACAGATGTCATTTCTAAGAACTAAAATTCTTTGGGAAAAACCCTCAAAAAA
AAAAAAAAAAAAAAAAAAAAA

486/615

FIGURE 482

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92289
><subunit 1 of 1, 234 aa, 1 stop
><MW: 26077, pI: 8.13, NX(S/T): 1
MAASRWARKAVVLLCASDLLLLLLLLLLPPPGSCAAEGSPGTPDESTPPPRKKKKDIRDYND
ADMARLLEQWEKDDDDIEEGDLPEHKRPSAPVDFSKIDPSKPESILKMTKKGKTLMMFVTV
SGSPTEKETEEITSLWQGSLFNANYDVQRFIVGSDRAIFMLRDGSYAWWEIKDFLVGQDRC
ADVTLEGQVYPGKGGGSKEKNKTKQDKGKKKKEGDLKSRSSKEENRAGNKREDL

Important features of the protein:

Signal peptide:

Amino acids 1-32

N-glycosylation site:

Amino acids 201-205

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 85-89

Tyrosine kinase phosphorylation site:

Amino acids 50-59

N-myristoylation sites:

Amino acids 30-36;138-144;153-159;176-182

Amidation site:

Amino acids 207-211

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

488/615

FIGURE 484

MALPPGPAALRHTLLLLLPALLSSGWGELEPQIDGQTAERALRENERHAFTCRVAGGPGTPRL
AWYLDGQLQEASTSRLLSVGGEAFSGGTSTFTVTAHRAQHNLCSLQDPRSGRSANASVILNV
QFKPEIAQVGAKYQEAQGPGLLVVLFALVRANPPANVTWIDQDGPVTVNTSDFLVLDAQNYPW
LTNHTVQLQLRSLAHNLSVVATNDVGVTASALPAPGPSRHPSLISSDSNNLKLNNVRLPRENM
SLPSNLQLNDLTPDSRAVKPADRQMAQNNSRPELLDPEPGGLLTSQGFIRLPVLGYIYRVSSV
SSDEIWL

N-glycosylation sites:

amino acids 106-110, 119-123, 162-166, 175-179, 192-196, 205-209,
251-255, 280-284

Glycosaminoglycan attachment site:

amino acids 23-27

Casein kinase II phosphorylation sites:

amino acids 36-40, 108-112, 164-168, 282-286, 316-320

N-myristoylation sites:

amino acids 34-40, 89-95, 215-221, 292-298, 293-299

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

489/615

FIGURE 485

AGAGTTCCTTTTTCTAGGTCGATTAGGTTATACATTGTTGAAGTATAGTTTCGAGTTAGAATT
GGTCATTTTATTTTCAGTGTTTCACAGAAATCGAAGAAGACAGAAATGGCGCTTCTGTGGTGG
ATATCTACAGTAGCAATACTGTTGTTTACTTCGACGATTTTGGAACATACGTTGAAGCTGGT
GCCGCTAAGTCTAACGAAGAAGAGATTGTGAACAAAAGCGAATTTGGAAGATTTCCACGAGGG
TCGAGAAAGGATGCATCGGGGTGCCACAAGCCGGGCTACCCCTGTACCCCTCATTCTCGCTGC
CCTCCACCTCCCCATGTGCAGCGTCCTCGTCCTATTCTGCATGCTTAGTCTAACACCATCAGG
CTCGTTTATCTTTTCTGTCAATTGATCTCACCAGGAGCAAATCACTAGTGCGTGCTTCTGATTC
ACGTAACGTAGTATGTAAATAAATGTCAGTGATATTATGAATTGGTAAAACATTTCTGTTATC
TAAATAAACAGTGAAGTTTGTGTTGACTAAAAAA

20241010 095250

490/615

FIGURE 486

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96855
><subunit 1 of 1, 84 aa, 1 stop
><MW: 9274, pI: 9.70, NX(S/T): 1
MALLWWISTVAILLFTSTILGTYVEAGAAKSNEEEIVNKSEFGRFPRGSRKDASGCHKPG
YPVPPHSRCPPPPHVQRP RPILHA
```

Signal peptide:

Amino acids 1-21

N-glycosylation site:

Amino acids 38-42

N-myristoylation site:

Amino acids 27-33

1005556 04130

CGGGGACGGAAGCGGCCCTGGGCCCGAGGGGCTGGAGCCGGGCCGGGGCG**ATGT**GAGCGCG
GGCCGCGGGGCTGCCCTGGCCGGTGCTGTTGGGGCTGCTGCTGGCGCTGTTAGTGCCGGGC
GGTGTTGCCGCCAAGACCGGTGCGGAGCTCGTGACCTGCGGGTCGGTGCTGAAGCTGCTCAAT
ACGCACCACCGCGTGCGGCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCAGCAA
TCGGTGACCGGCGTAGAGGCGTCGGACGACGCCAATAGCTACTGGCGGATCCGCGGCGGCTCG
GAGGGCGGGTGCCCGCGCGGGTCCCCGGTGCGCTGCGGGCAGGCGGTGAGGCTCACGCATGTG
CTTACGGGCAAGAACCTGCACACGCACCACTTCCCCGTGCGCGCTGTCCAACAACCAGGAGGTG
AGTGCCTTTGGGGAAGACGGCGAGGGCGACGACCTGGACCTATGGACAGTGCGCTGCTCTGGA
CAGCACTGGGAGCGTGAGGCTGCTGTGCGCTTCCAGCATGTGGGCACCTCTGTGTTCTGTCA
GTCACGGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGGCATGCCCAGT
GCCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTTCATCAAGCCTAGTGTGGAGCCC
TCTGCAGGTCACGATGAACT**CTGA**GTGTGTGGATGGATGGGTGGATGGAGGGTGGCAGGTGGG
GCGTCTGCAGGGCCACTCTTGGCAGAGACTTTGGGTTTGTAGGGGTCCTCAAGTGCCTTTGTG
ATTAAAGAATGTTGGTCTATGAAA

492/615

FIGURE 488

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96857
><subunit 1 of 1, 221 aa, 1 stop
><MW: 23598, pI: 6.96, NX(S/T): 0
MWSAGRGGAAWPVLLGLLLALLVPGGGAAKTGAELVTCGSVLKLLNTHHRVRLHSHDIKYGSG
SGQQSVTGVEASDDANSYWRI RGGSEG GCPRGSPVRCGQAVRLTHVLTGKNLHTHHFPSPLSN
NQEVS AFGEDGE GDDLDLWTVRC SGQHWEREA A VRFQHVGT SVFLSVTGEQY GSPIRGQHEVH
GMPSANTHNTWKAMEGIFIKPSVEPSAGHDEL

Important features of the protein:

Signal peptide:

amino acids 1-28

Glycosaminoglycan attachment site.

amino acids 62-66

N-myristoylation sites.

amino acids 16-22, 25-31, 27-33, 61-67, 71-77, 86-92, 87-93,
91-97, 190-196

Endoplasmic reticulum targeting sequence.

amino acids 218-223

205336 01100
205336 01100

FIGURE 489

CAGCAGCCGAGACAGCAGCTGAGACGGCAGCGGCAGCTTCTCAGGGCCGGAGCCAGTTCTTGAGGAGACTCTGC
 ACAGGGC**ATGG**ATCACTGTGGTGCCCTTTTCTGTGCCTGTGCCTTCTGACTTTGCAGAATGCAACAACAGAGAC
 ATGGGAAGAACTCCTGAGCTACATGGAGAATATGCAGGTGTCCAGGGGCCGGAGCTCAGTTTTTTCTCTCGTCA
 ACTCCACCAGCTGGAGCAGATGCTACTGAACACCAGCTTCCAGGGCTACAACCTGACCTTGACAGACACCCACCAT
 CCAGTCTCTGGCCTTCAAGCTGAGCTGTGACTTCTCTGGCCTCTCGCTGACCACTGCTGACTCTGAAGCGGGTGCC
 CCAGGCAGGAGGTGAGCATGCCCCGGGGTCAGCACGCCATGCAGTTCCCCGCCGAGCTGACCCGGGACGCCTGCAA
 GACCCGCCCCAGGGAGCTGCGGGCTCATCTGTATCTACTTCTCCAACACCCACTTTTTCAAGGATGAAAACAACTC
 ATCTCTGCTGAATAACTACGTCTGGGGGCCAGCTGAGTCATGGGCACGTGAACAACCTCAGGGATCCTGTGAA
 CATCAGCTTCTGGCACAAACCAAGCCTGGAAGGTACACCCTGACCTGTGTCTTCTGGAAGGAGGGAGCCAGGAA
 ACAGCCCTGGGGGGCTGGAGCCCTGAGGGCTGTGCTACAGAGCAGCCCTCCCACTCTCAGGTGCTCTGCCGCTG
 CAACCACCTCACCTACTTTGCTGTTCTCATGCAACTCTCCCCAGCCCTGGTCCCTGCAGAGTTGCTGGCACCTCT
 TACGTACATCTCCCTCGTGGGCTGCAGCATCTCCATCGTGGCCTCGCTGATCACAGTCTGCTGCACTTCCATTT
 CAGGAAGCAGAGTGACTCCTTAACACGTATCCACATGAACCTGCATGCCCTCCGTGCTGCTCCTGAACATCGCCTT
 CCTGCTGAGCCCCGATTCGCAATGTCTCCTGTGCCCGGGTCAGCATGCACGGCTCTGGCCGCTGCCCTGCACTA
 CGCTGGCTCAGCTGCCCTCACCTGGATGGCCATCGAGGGCTTCAACCTCTACCTCCTCCTCGGGCGTGTCTACAA
 CATCTACATCCGCAGATATGTGTTCAAGCTTGGTGTGCTAGGCTGGGGGGCCCCAGCCCTCCTGGTGCTGCTTTT
 CCTCTCTGTCAAGAGCTCGGTATACGGACCCCTGCACAATCCCCGTCTTCGACAGCTGGGAGAATGGCACAGGCTT
 CCAGAACATGTCCATATGCTGGGTGCGGAGCCCCGTGGTGCACAGTGTCTGGTTCATGGGCTACGGCGGCCCTCAC
 GTCCCTCTTCAACCTGGTGGTGTGCTGGCCTGGGCGCTGTGGACCTGCGCAGGCTGCGGGAGCGGGCGGATGCACC
 AAGTGTGAGGGCTGCCATGACACTGTCACTGTGCTGGGCTCACCGTGTGCTGGGAACCACTGGGCTTGGC
 CTTCTTTTCTTTTGGCGTCTTCTGCTGCCCCAGCTGTCTCTCTTACCCTCTTAACTCGCTGTACGGTTTCTT
 CCTTTTCTGTGGTTCTGCTCCAGCGGTGCCGCTCAGAAGCAGAGGCCAAGGCACAGATAGAGGCCCTTCAGCTC
 CTCCCAAACAACACAG**TAG**TCCGGGCTCCTGGCCTGGAATCCTCAGCCTCTCTGGCCGCCAGTAGCCTGAGGCT
 ACGGCTCCTGCTAGAGAGGTGGCAGGCCTGCTGCTGGACCCAGAGGCCACTGTGACCGCCAAGGGGCTTTTCT
 CACTTCCACGGCTCTCCAGGCACTGAGGGGAAGGCATTGCTCTACCTCTCCCTGACATTTTGCTCCGGGGCAGA
 TCCAACCTTACCTGGGGCAGCAAACTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 ACAGACCTTAGGTATCCACAGCTGTGACATGGGGGCAAGCAGCTTTGTCTTACGCTAACCAGGAGCTTAGTAAA
 AATTGCATAAGACCAGGGGGAAGAGTGTGAGCGTGGGGTGGGAATTCCCGCGGCTCCACCTGCTTGTAGGGGC
 AGGATCTCATTCAGGCTGCCCTGGAAGCACCTGCTTGGCCCTGCCACCTTCTCCAGGGGAGGGCCAGATGGCAT
 CCTGGCTTGGGGCGGGTGGGACCTACCCAGGCTCTGAGACTTTACTGGCCTATGCCTGAGGCCTCTTTTCTTTTA
 ACTCCCTAAATTTATGATGACTCCAAGTCCAAGCCACCTTCCCAAAGATTGGGAGGTTCGCCGCTTCCAGAGG
 CTTCTCTGCGGTGCTCCCAAGACTTCCATAGACCATCTGGACCACTAGCCCATCCCGCAGTTTTCTTGGGGGCA
 GAGGAAAACGCTTCTTTCTCCTCCAGCTGAATCAGCTGGATCCAGTGTCTGGCTGTTTGGTGAATTGGGCAAGA
 TTGAATTTGCCAGGTAGGCGTGAGAGTGTGGGTTTTAAATTCGAAGCTCAGGCCATAGTTTCAGAGAATCACCC
 TTACCCAGACCTTCATGAGACAGTGTCTCATGAAGCCAGTGCCTTTCCAGAACGAACACTAGGCGGCACCGTTG
 GTCCACACTCAGAGGCCCTTGGCGCCAAGACTGCATCTAGAATCGCTCAAACACCTGTTTGAGACCCCATGCAC
 CAGCTGGAGGGGCCGTAACCTGCAGGACTGCGCCTACTGAGTGACCCATTTCTTCCAGGAGGAAAGGCAAGACAG
 CTTACACGGCCATTTGTCTCTTTTCCCAATGCGGCGGTGCACTTTTCGCTCTTGGGGGCTGCACCCAGACATAGC
 TGGCACCAAGAGCAGGGTGTCTCAGGTGGTGGGTGCTCAGGGCCCTGCCAGGCCACTGGGCGCTTTTGATGACCT
 CAAAGGTACAGGCAGAAAATAGGAGCAGGATTTCCCTGGGAAAAGTTATCCTGGGACATCTTCTGCTCTTCT
 GTACATTTCTAGATGCAATAACTCCTTACCAGGCAGTGAGTGGCGTAGGCTCTGGAGCCAGGCTGCCTGGGCT
 CCAATGCCAGCTCTGCCACTTGCTAGCTGTGAGACTGTGGACAAACCACTCAGCCTCTGTGTGCTCAGTTTTCC
 TATTTGTAAATAGAGACCATAGTGGTACCTATTTTGAAGACTAAGTAAAAGAATTCAAATAAAGAGACTTGGCA
 CAGAGTAAGTGCTCAGTAAAAA

494/615

FIGURE 490

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96860
><subunit 1 of 1, 528 aa, 1 stop
><MW: 59000, pI: 8.73, NX(S/T): 9
MDHCGALFLCLCLLTQNATTETWEELLSYEMNQVSRGRSSVFSSRQLHQLEQMLLNTS
FPGYNLTLOTPTIQSLAFKLSCDFSGLSLT SATLKRVPQAGGQHARGQHAMQFPAELTRD
ACKTRPRELRLICIYFSNTHFFKDENNSSLNLYVLGAQLSHGHVNNLRDPVNISFWHNQ
SLEGYTLTCVFWKEGARKQPWGGWSPEGCRTEQPSHSQVLCRCNHLTYFAVLMQLSPALV
PAELLAPLTYISLVGCSISIVASLITVLLHFHFRKQSDSLTRIHMNLHASVLLLNIAFLL
SPAFAMSPVPGSACTALAAALHYALLSCLTWMAIEGFNLYLLLGRVYNIYIRRYVFKLGV
LGWGAPALLVLLSLSVKSSVYGPTIPVFD SWENGTGFQNM SICWVRSPVHVSVLVMGYG
GLTSLFNLVVLAWALWTLRRLRERADAPSVRACHDTVTVLGLTVLLGTTWALAFFSFGVF
LLPQLFLFTILNSLYGFFLFLWFCSQRCRSEAEAKAQIEAFSSSQTTQ

Important features of the protein:

Signal peptide:

Amino acids 1-21

Transmembrane domains:

Amino acids 244-264;290-309;316-344;358-376;411-431;468-491

N-glycosylation sites:

Amino acids 18-22;58-62;65-69;146-150;147-151;173-177;
179-183;394-398;400-404

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 274-278

N-myristoylation sites:

86 GLSLTS
101 GGQHAR
157 GAQLSH
255 GCSISI
311 GSACTA
420 GGLTSL
467 GTTWAL

Prokaryotic membrane lipoprotein lipid attachment sites:

Amino acids 246-257;318-329

Eukaryotic thiol (cysteine) proteases histidine active site:

Amino acids 410-421

G-protein coupled receptors family 2 proteins:

Amino acids 273-302;314-343

2003336.0140

FIGURE 491

CTTGGCTGCCCCGACAACAAGCTCGCCACCTGCGCTGGGCGCATCCACCATCCAAGGCCAGCT
GAGGGGCACCAGACAGAGGATGAGGAGAGAGAGTCCGACACGGGCTGCCCTGAGAGACATTTTC
CATGGACATCCTCATGCTGCTTCTGCTTTTGTGTGTAATATATGGGAGATTTTCCCAAGATGA
ATACTCCCTCAATCAAGCTATCCGGAAAGAATTTACAAGAAATGCCAGAAACTGCTTGGGTGG
CCTGAGAAACATCGCTGACTGGTGGGACTGGAGTCTGACCACACTTCTGGATGGCCTGTACCC
GGGAGGCACCCCGTCAGCCCGTGTGCCGGGGGCTCAGCCTGGAGCTCTTGGAGGAAATGCTA
CCTAATAGGCAGTTCCGTAATTAGGCAGCTAAAAGTTTTTCTAGGCATTTATGCAAGCCTCC
CAGGCCATTTTCAGCACTCATCGAAGACTCTATTCTACATGTAGTCCCGAAGTTGGAGGCCC
TGAGAACCCCTACCTGATAGACCCAGAGAACCAAAACGTGACCCTGAATGGTCCTGGGGGCTG
TGGGACAAGGGAGGACTGTGTGCTCAGCCTGGGCAGAACAAAGGACTGAAGCCCACACAGCCCT
GTCCCGACTCAGGGCCAGCATGTGGATTGACCGCAGCACCAGGGCTGTGTCTGTGCACTTCAC
TCTCTATAACCCTCCAACCCAACCTCTTACCAGCGTGTCCCTGAGAGTGGAGATCCTCCCTAC
GGGGAGTCTCGTCCCTCATCCCTGGTGGAGTCATTGAGCATCTTCCGCAGCGACTCAGCCCT
GCAGTACCACCTCATGCTTCCCCAGCTGGTCTTCCCTGGCACTCAGCCTGATCCACCTCTGTGT
TCAACTCTACCGTATGATGGACAAGGGCGTCTCAGCTACTGGCGAAAGCCAAGGAAGTGGCT
GGAGGTAGCCTCTCTTGTGTGTCATTTTCTTTTGAAAATAACAATAAACTGTTTATATCTTGAA
AAAATAATTTAAATAAGAAATTGATTATGCACTAGCTACTGCCAACATTATTGCAGTTTCTC
CCTCTGTAGTGTTAATCTCAAAACAGCATTTGAGATCAGGTATCATTTAGTGTTGTTACAGTT
ACCGTCATGTACCACACGAATTTAGCCAAGGTGGTGGTCCATAAGATCATATGGTGCTAAG
AAATTTCTGTCACCTAATGACATCTTGATTCTGACCTTGTATGTAGGCCTAGGCTAAATATGT
CTGTTTGTATCTTAGCTTTTAATAAAGAAGTTTAAAAATAAAAAA

496/615

FIGURE 492

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96861
><subunit 1 of 1, 300 aa, 1 stop
><MW: 33649, pI: 9.26, NX(S/T): 1
MRRESRTRAALRDISMDILMLLLLLLCVIYGRFSQDEYSLNQAIRKEFTRNARNCLGGLRN
IADWWDWSLTTLDDGLYPGGTPSARVPGAQPGALGGKCYLIGSSVIRQLKVFPRLCKPP
RPFSAIEDSIPTCSPEVGGPENPYLIDPENQNVTLNGPGGCGTREDCVLSLGRTRTEAH
TALSRLRASMWIDRSTRAVSVHFTLYNPPTQLFTSVSLRVEILPTGSLVPSSLVESFSIF
RSDSALQYHMLPQLVFLALSLIHLVCVQLYRMMDKGVLSYWRKPRNWLEVASLVSFSEK
```

Important features of the protein:

Signal peptide:

Amino acids 1-30

Transmembrane domain:

Amino acids 250-267

N-glycosylation site:

Amino acids 153-157

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 2-6

N-myristoylation sites:

Amino acids 56-62;75-81;79-85;80-86;88-94;92-98;160-166

496/615

FIGURE 493

TCTCAGGGCTTCATACAGGAAATCTATTGCTGTGTCAAGTTCCAGAGAAAAGCTTCTGTTCGT
 CCAAGTTACTAACCAGGCTAAACCACATAGACGTGAAGGAAGGGGCTAGAAGGAAGGGAGTGC
 CCCACTGTTGATGGGGTAAGAGGATCCTGTACTGAGAAGTTGACCAGAGAGGGTCTCACCATG
 CGCACAGTTCCCTTCTGTACCTGTGTGGAGGAAAAGTACTGAGTGAAGGGCAGAAAAAGAGAAA
 ACAGAAATGCTCTGCCCTTGGAGAACTGCTAACCTAGGGGCTACTGTTGATTTTGA CTATCTTC
 TTAGTGGCCGAAGCGGAGGGTGCTGCTCAACCAACAACCTCATTAATGCTGCAAACTAGCAAG
 GAGAATCATGCTTTAGCTTCAAGCAGTTTATGTATGGATGAAAAACAGATTACACAGAACTAC
 TCGAAAGTACTCGCAGAAGTTAACACTTCATGGCCTGTAAAGATGGCTACAAATGCTGTGCTT
 TGTGCCCCTCCTATCGCATTAAAGAAATTTGATCATAATAACATGGGAAATAATCCTGAGAGGC
 CAGCCTTCCTGCACAAAAGCCTACAGGAAAGAAACAAATGAGACCAAGGAAACCAACTGTACT
 GATGAGAGAATAACCTGGGTCTCCAGACCTGATCAGAATTCGGACCTTCAGATTTCGTCCAGTG
 GCCATCACTCATGACGGGTATTACAGATGCATAATGGTAACACCTGATGGGAATTTCCATCGT
 GGATATCACCTCCAAGTGTTAGTTACACCTGAACTGACCCTGTTTCAAACAGGAATAGA ACT
 GCAGTATGCAAGGCAGTTGCAGGGAAGCCAGCTGCGCAGATCTCCTGGATCCCAGAGGGCGAT
 TGTGCCACTAAGCAAGAATACTGGAGCAATGGCACAGTGA CTGTTAAGAGTACATGCCACTGG
 GAGGTCCACAATGTGTCTACCGTGACCTGCCACGTCTCCCATTTGACTGGCAACAAGAGTCTG
 TACATAGAGCTACTTCCTGTTCCAGGTGCCAAAAAATCAGCAAAATTATATATTCATATATC
 ATCCTTACTATTATTATTTTTGACCATCGTGGGATTTCATTTGGTTGTTGAAAGTCAATGGCTGC
 AGAAAAATATAAATTGAATAAAACAGAATCTACTCCAGTTGTTGAGGAGGATGAAATGCAGCCC
 TATGCCAGCTACACAGAGAAGAACAATCCTCTCTATGATACTACAAACAAGGTGAAGGCATCT
 CAGGCATTACAAAGTGAAGTTGACACAGACCTCCATACTTTATAAGTTGTTGGACTCTAGTAC
 CAAGAAACAACAACAAACGAGATACATTATAATTACTGTCTGATTTTCTTACAGTTCTAGAAT
 GAAGACTTATATTGAAATTAGGTTTTCCAAGGTTCTTAGAAGACATTTTAATGGATTCTCATT
 CATACCCTTGATATAATTGGAATTTTTTGATTCTTAGCTGCTACCAGCTAGTTCTCTGAAGAACT
 GATGTTATTACAAAGAAAATACATGCCCATGACCAAATATTCAAATTGTGCAGGACAGTAAAT
 AATGAAAACCAAATTTCTCAAGAAATAACTGAAGAAGGAGCAAGTGTGAACAGTTTCTTGTG
 TATCCTTTTCAGAATATTTTAATGTACATATGACATGTGTATATGCCTATGGTATATGTGTCAA
 TTTATGTGTCCCCTTACATATACATGCACATATCTTTGTCAAGGCACCAGTGGGAACAATACA
 CTGCATTACTGTTCTATACATATGAAAACCTAATAATATAAGTCTTAGAGATCATTTTATATC
 ATGACAAGTAGAGCTACCTCATTCTTTTTAATGGTTATATAAAATTCCATTGTATAGTTATAT
 CATTATTTAATTAAAAACAACCCTAATGATGGATATTTAGATTCTTTTAAGTTTTGTTTATTT
 CTTTTAAGTTTTGTTTGTGGTATAAACAATACCACATAGAATGTTTCTTGTTTCATATATCTCT
 TTGTTTTTGAGTATATCTGTAGGATAACTTTCTTGAGTGGAATTGTCAGGTCAAAGGGTTTGT
 GCATTTTACTATTGATATATATGTTAAATTGTGTCAAATATATATGTCAAATTCCTCCAACA
 TTGTTTAAATGTGCCTTTCCCTAAATTTCTATTTTAATAACTGTACTATTCTGCTTCTACAG
 TTGCCACTTTCTCTTTTAAATCAACCAGATTAAATATGATGTGAGATTATAATAAGAATTATA
 CTATTTAATAAAAATGGATTTATA

CCAGTGTGCACACGCGCATTCGCCCGAGGCTGTCAACCGCCCTGCCCGCCACCCAGCTGTCTTG
GACCCAGGGGCAGGGAGAGGCTGGACGCCAGGTGCGCGGACACAGAAGCGTCTAAGCACAGCT
TCCTCCTTGCCGCTCCGGGAAGTGGGCAGCCAGCCAGGAACCAGTACCACCTGCACCATGGG
GCTGTCCCGGAAGGAGCAGGTCTTTCTTGCCCTGCTGGGGGCCTCGGGGGTCTCAGGCCCTCAC
GGCACTCATTTCTCCTCCTGGTGGAGGCCACCAGCGTGCTCCTGCCACAGACATCAAGTTTG
GATCGTGTTTTGATGCGGGCTCCTCCACACGTCCCTCTTCTGTATCAGTGCCGGCGAACA
GGAGAATGGCACGGGTGTGGTCAGCCAGGCCCTGGCCTGCCAGGTGGAAGGGCCTGGAATCTC
CTCCTACACTTCTAATGCTGCACAGGCTGGTGAGAGCCTGCAGGGCTGCTTGAGGAGGCGCT
GGTGCTGATCCAGAGGCCCAGCATCGGAAAACACCCACGTTCTTGGGGGCCACGGCTGGCAT
GAGGTTGCTCAGCCGGAAGAACAGCTCTCAGGCCAGGGACATCTTTGCAGCAGTCACCCAGGT
CCTGGGCCCGTCTCCCGTGGACTTTTGGGGTGCCGAGCTCCTGGCCGGGCAGGCCGAAGGTGC
CTTTGGTTTGATCACTGTCAACTACGGCTTGGGGACGCTGGTCAAGTACTCCTTCACTGGAGA
ATGGATCCAGCCTCCGGAGGAGATGCTGGTGGGTGCCCTGGACATGGGAGGGGCCTCCACCCA
GATCACGTTTCGTGCCTGGGGGCCCATCTTGACAAGAGCACCCAGGCCGATTTTCGCCTCTA
CGGCTCCGACTACAGCGTCTACACTCACAGCTACCTGTGCTTTGGACGGGACCAGATGCTGAG
CAGGCTCCTCGTGGGGCTGGTGCAGAGCCGCCCGGCTGCCCTGCTCCGTACCCGTGCTACCT
CAGCGGCTACCAGACACACTGGCCCTGGGCCCGCTGTATGAGTCACCCTGTGTCCACGCCAC
GCCCCGCTGAGCCTCCCCCAGAACCTCACAGTTGAAGGGACAGGCAACCCTGGAGCCTGCGT
CTCAGCCATCCGGGAACTTTTCAACTTCTCCAGCTGCCAGGGCCAGGAGGACTGCGCCTTTGA
CGGGTCTACCAGCCCCCGCTGCGGGGCCAGTTCTATGTGGAGGCCAGCTACCCTGGGCAGGA
CCGCTGGCTGCGGGACTACTGTGCCTCAGGCCTGTACATCCTCACCCCTCCTGCACGAGGGCTAC
GGGTTCAGCGAGGAGACCTGGCCAGCCTCGAGTTCCGAAAGCAGGCGGGCGGTGTGGACATT
GGCTGGACACTGGGCTACATGCTGAACCTGACCGGGATGATCCCGGCCGATGCGCCGGCTCAG
TGGCGGGCAGAGAGCTACGGCGTCTGGGTGGCCAAAGTGGTGTTTATGGTGCTGGCCCTGGTG
GCGGTGGTGGGGGCTGCCTTGGTCCAGCTCTTCTGGTTGCAGGACTTAGTGGGAAGGCGGAGGT
GGGCCCCCACAGAGCCACAGGCAGCTGCGTCCCGGATGCTGGAGGCTTCTGAGCCCTGAGC
GCCGTGGGGCCTTGCTCTGTGGCTCTGCCACGGTCAGGTGACAGCCACCTCCAGGGCACCGT
CAGGGTGGTGCTGGCCACAGAGGCTGCATGACCTCCCCTCCCGGCGTCCCTCCCCAACCTCC
TTCCGCAACTGGGCTTCCAGGGCCGTAGGTGCCTTTCTGCACACAGGCCGCCAGGACTCGTG
TGCTCTCCAGGCTGTGTGACTGCAGGGCCACATGCTGCCTGCAAACAGGGCAAGACCACGGAG
CACAGGGGTCTTGCTCCTGATGGGGCCTCAGGAGGGGCGGAGAGGGGTGGAAGGGAGGGAGCT
GCCCCACCTGGACCCCCGCTCTCCCTGCTGTTGTCTGAGCAGATGGATGGAGTCCAGGCCTGG
GGGCTTCTGCTGGGCCAGCCGGCCTCCACACCCACTTGAGGGGTGAGACTGCAGTGGGGGT
TGTTTTTTATTAAAAAGCATCATGGACACAGCAAAAAAAAAAAAAAAAAA

500/615

FIGURE 496

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96870
><subunit 1 of 1, 458 aa, 1 stop
><MW: 49377, pI: 4.98, NX(S/T): 5
MGLSRKEQVFLALLGASGVSGLTALILLVLEATSVLLPTDIKFGIVFDAGSSHTSLFLYQ
WPANKENGTGVVSQALACQVEGPGISSYTSNAAQAGESLQGCLEEALVLIPEAQHRKTPT
FLGATAGMRLLSRKNSSQARDIFAAVTQVLGRSPVDFWGAELLAGQAEGAFGWITVNYGL
GTLVKYSFTGEWIQPPEEMLVGALDMGGASTQITFVPGGPILDKSTQADFRLYGSDYSVY
THSYLCFGRDQMLSRLLVGLVQSRPAALLRHPCYLSGYQTTLALGPLYESPCVHATPPLS
LPQNLTVEGTGNPGACVSAIRELFNFSSCQGQEDCAFDGVYQPPLRGQFYVEASYPGQDR
WLRDYCASGLYILTLLHEGYGFSEETWPSLEFRKQAGGVDIGWTLGYMLNLTGMI PADAP
AQWRAESYGVVWAKVVMVLALVAVVGAALVQLFWLQD

Important features of the protein:

Signal peptide:

Amino acids 1-21

Transmembrane domain:

Amino acids 428-449

N-glycosylation sites:

Amino acids 67-71;135-139;304-308;325-329;410-414

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 133-137

N-myristoylation sites:

Amino acids 50-56;123-127;165-171;207-213;234-240;
259-265;311-317;314-320;331-337;398-404;
413-419;429-435

GDA1/CD39 family of nucleoside phosphatases proteins:

Amino acids 43-59;202-215

FIGURE 497

GCCTTATAAAGTAGCCTCTGCATCTGCCTGCCTCGGGCAGAGGAGGGCTACCCTGGGGCTGAG
 AGTTCACCTGTCTCAGGAACCACCTGAGCCACAGATCCTGTGGGCAGCGGCCAGGGCAGCCA
TGGCTTGGGCAAGTAGGCTGGGCCTGCTGCTGGCACTGCTGCTGCCCCGTGGTTCGGTGCCTCCA
 CGCCAGGCACCGTGGTCCGACTCAACAAGGCAGCATTGAGCTACGTGTCTGAAATTGGGAAAG
 CCCCTCTCCAGCGGGCCCTGCAGGTCACTGTCCCTCATTTCTGGACTGGAGTGGAGAGGCGC
 TTCAGCCCACCAGGATCCGGATTCTGAATGTCCATGTGCCCCGCCTCCACCTGAAATTCATTG
 CTGGTTTTCGGAGTGCGCCTGCTGGCAGCAGCTAATTTTACTTTCAAGGTCTTTCGCGCCCCAG
 AGCCCCCTGGAGCTGACGCTGCCTGTGGAAGTCTGGCTGACACCCGCGTGACCCAGAGCTCCA
 TCAGGACCCCTGTGGTCAGCATCTCTGCCTGCTCTTTATTCTCGGGCCACGCCAACGAGTTTG
 ATGGCAGTAACAGCACCTCCCACGCGCTGCTGGTCTGCTGGTGCAGAAGCACATTAAAGCTGTCT
 TGAGTAACAAGCTGTGCCTGAGCATCTCCAACCTGGTGCAGGGTGTCAATGTCCACCTGGGCA
 CCTTAATTGGCCTCAACCCCGTGGGTCTGAGTCCCAGATCCGCTATTCCATGGTCAGTGTGC
 CCACTGTCAACAGTGACTACATTTCCCTGGAAGTCAATGCTGTTCTCTTCCTGCTGGGCAACC
 CCATCATCCTGCCCACGGATGCCACCCCTTTTGTGTTGCCAAGGCATGTGGGTACCGAGGGCT
 CCATGGCCACCGTGGGCCTCTCCCAGCAGCTGTTTGACTCTGCGCTCCTGCTGCTGCAGAAGG
 CCGGTGCCCTCAACCTGGACATCACAGGGCAGCTGAGGTCCGATGACAACCTGCTGAACACCT
 CTGCTCTGGGCCCGGCTCATCCCAGAGGTGGCCCGCCAGTTTCCCAGCCCATGCCTGTGGTGC
 TCAAGGTGCGGCTGGGTGCCACACCTGTGGCCATGCTCCACACAAACAACGCCACCCCTGCGGC
 TGCAGCCCTTCGTGGAGGTCTTGCCACAGCCTCCAACCTCGGCTTTCCAGTCCCTCTTCTCCC
 TGGATGTGGTAGTGAACTTGAGACTCCAGCTCTCTGTGTCCAAGGTGAAGCTTCAGGGGACCA
 CGTCTGTGCTGGGGGATGTCCAGCTCACGGTGGCCTCCTCCAACGTGGGCTTCATTGATACAGAT
 CAGGTGCGCACACTGATGGGCACCGTTTTTGGAGAAGCCCCTGCTGGACCATCTCAATGCTCTC
 TTGGCCATGGGAATTGCCCTCCCTGGTGTGGTCAACCTCCACTATGTTGCCCTGAGATCTTT
 GTCTATGAGGGCTACGTGGTGATATCCAGTGGACTCTTCTACCAGAGCTGAGGCAAGACCACT
 GGGAGGCCTGAGAGTGGGCCAGCTCGCTGCTCAGGCGAATTTCTCATTTCAAGCCACTGGGGA
 AACTGAGGCAAAACCATACTTAGTCATCACCAACAAGCTGGACTGCTTAGCTGGGCTGTTTTA
 TCTTCCCTGAGTGCCTGGGTCTCCCTCCCTCACTTCTGCCCTTTCCCTTCCTCCTCCTCTTCT
 CCTCCCTCTTCCCTCATCTCCCCCTCCTTCCTCTGCCCCACCCAGGGGGGAGCAGACTGCT
 CCTCCAGGCTGTATAGACCTGCCCTCTTGCAATTAAACAACCTTCTCTTGAGCTGC

502/615

FIGURE 498

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96872
><subunit 1 of 1, 458 aa, 1 stop
><MW: 49158, pI: 8.72, NX(S/T): 4
MAWASRLGLLLALLLPVVGASTPGTVVRLNKAALSYVSEIGKAPLQRALQVTVPHFLDWS
GEALQPTRIRILNVHVPRLHLKFIAGFGVRLLAANFTFKVFRAPPELELTLPVELLADT
RVTQSSIRTPVVSISACSLFSGHANEFDGSNSTSHALLVLVQKHIKAVLSNKLCLISISNL
VQGVNVHLGTLIGLNPVGPESQIRYSMVSVPTVTSYISLEVNAVLFLLGNPIILPTDAT
PFVLPRHVGTEGSMATVGLSQQLFDSALLLLQKAGALNLDITGQLRSDDNLLNTSALGRL
IPEVARQFPEPMPVVLKVRLGATPVAMLHTNNATLRLQPFVEVLATASNSAFQSLFSLDV
VVLNRLQLSVSKVKLQGTTSVLGDVQLTVASSNVGFIDTDQVRTLMGTVFEKPLLDHLNA
LLAMGIALPGVVNLHYVAPEIFVYEGYVVISSGLFYQS

Important features of the protein:

Signal peptide:

Amino acids 1-20

Transmembrane domain:

Amino acids 217-236

N-glycosylation sites:

Amino acids 96-100;151-155;293-297;332-336

N-myristoylation sites:

Amino acids 8-14;149-155;189-195;249-255;252-258;283-289

LBP / BPI / CETP family proteins:

Amino acids 22-50; 251-287

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

503/615

FIGURE 499

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCC**ATG**AAAGC
CCTTATGCTGCTCACCCCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTGCTGTCACTC
CTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGAGCCAGGACA
GCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAATCTGCGCTGTGG
CACACCAGAAGAGCCCTGTGAGGAGCCCTTCAACCAAACCAACCGCAAGCTGGGTCTGACATA
TAACACCACCTGCTGCAACAAGGACAACCTGCAACAGCGCAGGACCCCGGCCCACTCCAGCCCT
GGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTGGCTGCTGCAC**TGAG**ACTCATT
CCATTGGCTGCCCCCTCCTCCACCTGCCTTGGCCTGAGCCTCTCTCCCTGTGTCTCTGTATCC
CCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCATTTCTTTAATTAAACACTGTTCCGAGTG
GTCTCCTCATCCATCCTTCCACCTCACACCCCTTCACTCTCCTTTTTCTGGGTCCCTTCCAC
TTCCTTCCAGGACCTCCATTGGCTCCTAGAAGGGCTCCCCACTTTGCTTCCTATACTCTGCTG
TCCCTACTTGAGGAGGGATTGGGATCTGGGCCTGAAATGGGGCTTCTGTGTTGTCCCCAGTG
AAGGCTCCCACAAGGACCTGATGACCTCACTGTACAGAGCTGACTCCCCAAACCCAGGCTCCC
ATATGTACCCCATCCCCATACTCACCTCTTCCATTTTGAGTAATAAATGTCTGAGTCTGGA
AAAAAAAAAAAAAAAAAAAA

504/615

FIGURE 500

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96878
><subunit 1 of 1, 125 aa, 1 stop
><MW: 13821, pI: 8.60, NX(S/T): 2
MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDRQSCRLEPGQQCLTTHAYLGKMWVFSNL
RCGTPEEPCQEAFNQTNRKLGITYNTTCCNKDNCNSAGPRPTPALGLVFLTSLAGLGLWLLH
```

Important features of the protein:

Signal peptide:

amino acids 1-18

N-glycosylation sites.

amino acids 77-81, 88-92

N-myristoylation site.

amino acids 84-90

Ly-6 / u-PAR domain protein signature.

amino acids 85-98

504/615

505/615

FIGURE 501

GGAGCCTCCTAATGCAGTCTTCTGCACAGTCCTGGGGACTGACTGACTGAATCACACCTCTGG
GGCTGGGGGCTGCTGACATGTGTGCCTTTTCCTTGGCTGCTTCTTCTCCTGCTGCTCCAGGAGG
GCAGCCAAAGGAGACTCTGGAGATGGTGTGGATCCGAGGAAGTGGTTGCGGTCCTTCAGGAGT
CCATCAGCCTCCCCCTGGAAATACCACCAGATGAAGAGGTTGAGAACATCATCTGGTCCTCTCAC
AAAAGTCTTGCCACTGTGGTGCCAGGGAAAGAGGGACATCCAGCTACCATCATGGTGACCAAT
CCACACTACCAGGGCCAAGTGAGCTTCCTGGACCCCAGCTATTCCCTGCATATCAGCAATCTG
AGCTGGGAGGATTCAGGGCTTTACCAAGCTCAAGTCAACCTGAGAACATCCAGATCTCTACC
ATGCAGCAGTACAATCTATGTGTCTACCATCCTAACTATGCTTCTGAGAAGCCTTCAACAGCC
TTCTGCCTCCTGGCCAAGGGATTGCTCATCTTCTTGCTCTTGGTAATTCTGGCCATGGGACTC
TGGGTCATCCGAGTCCAGAAAAGACACAAAATGCCAAGGATGAAGAAACTCATGAGAAACAGA
ATGAAATTGAGGAAGGAGGCAAAGCCTGGCTCCAGCCCTGCCTGACTGCTCCTTGGGAACCCC
AGTCCTGAGCTTGCTTTCTTCCCAGCACCCAGAGAATCCTTCCTCAGCTCTCTTCTTCCAGG
GGAAGGAGGTGCTCAGGGGTGGGTATCCAGAGAGCCATACTTCTGAGGGAAGACTGGCTGGCA
ATAAAGTCAAATTAAGTGACCACA

505/615

506/615

FIGURE 502

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96879
><subunit 1 of 1, 198 aa, 1 stop
><MW: 22584, pI: 9.40, NX(S/T): 1
MCAFPWLLLLLLLLLQEGSQRRLLWRWCGSEEVVAVLQESISLPLEIPPDEEVENIIWSSHKS
LATVVPKGEGHPATIMVTNPHYQGQVSFLDPSYSLHISNLSWEDSGLYQAQVNLRTSQIS
TMQQYNLCVYHPNYASEKPSTAFCLLAKGLLIFLLLVLAMGLWVIRVQKRHKMPRMKKL
MRNRMKLRKEAKPGSSPA

Important features of the protein:

Signal peptide:

Amino acids 1-18

Transmembrane domain:

Amino acids 144-165

N-glycosylation site:

Amino acids 99-103

N-myristoylation site:

Amino acids 106-112

506/615

507/615

FIGURE 503

ACGGGCCGCGAGCGGCAGTGACGTAGGGTTGGCGCACGGATCCGTTGCGGCTGCAGCTCTGCAG
TCGGGCCGTTTCCTTCGCCGCCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCGCGCTAC
CGCACCCAGGTTTCGGCCCGTAGGCGTCTGGCAGCCCCGGCGCCATCTTCATCGAGCGCCATGGC
CGCAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCAATTTGTTTCTGCT
GACCGCGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCTGCGGGATGTAAAAGC
TCTTACCCTCCACTATGACCGCTATACCACCTCCCGCAGGCTGGATCCCATCCCACAGTTGAA
ATGTGTTGGAGGCACAGCTGGTTGTGATTCTTATACCCCAAAGTCATACAGTGTGAGAACAA
AGGCTGGGATGGGTATGATGTACAGTGGGAATGTAAGACGGACTTAGATATTGCATACAAATT
TGGA AAAACTGTGGTGAGCTGTGAAGGCTATGAGTCCTCTGAAGACCAGTATGTACTAAGAGG
TTCTTGTTGGCTTGAGTATAATTTAGATTATACAGAACTTGGCCTGCAGAACTGAAGGAGTC
TGGA AAGCAGCACGGCTTGCCTCTTTCTCTGATTATTATTATAAGTGGTCCTCGGCGGATTC
CTGTAACATGAGTGGATTGATTACCATCGTGGTACTCCTTGGGATCGCCTTTGTAGTCTATAA
GCTGTTCTTGAGTGACGGGCAGTATTCTCCTCCACCGTACTCTGAGTATCCTCCATTTTCCCA
CCGTTACCAGAGATTCACCAACTCAGCAGGACCTCCTCCCCCAGGCTTTAAGTCTGAGTTCAC
AGGACCACAGAATACTGGCCATGGTGCAACTTCTGGTTTTTGGCAGTGCTTTTACAGGACAACA
AGGATATGAAAATTCAGGACCAGGGTTCTGGACAGGCTTGGGAACTGGTGAATACTAGGATA
TTTGTTTTGGCAGCAATAGAGCGGCAACACCCTTCTCAGACTCGTGGTACTACCCGTCCTATCC
TCCCTCCTACCCTGGCACGTGGAATAGGGCTTACTCACCCCTTCATGGAGGCTCGGGCAGCTA
TTCGGTATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAGGATATGGTGGTACCAGGAG
ACGATTAAAGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAATTTTGGATTTTTTCATCACTTT
CTCTTTAGAAAAAAGTACTACCTGTAAACAATTGGGAAAAGGGGATATTCAAAGTTCTGTG
GTGTTATGTCCAGTGTAGCTTTTTGTATTCTATTATTTGAGGCTAAAAGTTGATGTGTGACAA
AATACTTATGTGTTGTATGTCAGTGTAAACATGCAGATGTATATTGCAGTTTTTGAAAGTGATC
ATTACTGTGGAATGCTAAAAATACATTAATTTCTAAAACCTGTGATGCCCTAAGAAGCATTAA
GAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAAAATTTAGTTTTAGGTGGTTGTAGC
TGATGAGTTATTACCTCATAGAGACTATAATATTCTATTTGGTATTATATTATTTGATGTTTG
CTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAATTATGCTAATTTGTGAGTTCTGATCAC
TTTTGAGCTCTGAAGCTTTGAATCATTCAGTGGTGGAGATGGCCTTCTGGTAACTGAATATTA
CCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTAGAAGGTTGTTGTGAATGACTCTGTGCTGG
CAAAAATGCTTGAAACCTCTATATTTCTTTCGTTTATAAGAGGTAAAGGTCAAATTTTTCAAC
AAAAGTCTTTTAATAACAAAAGCATGCAGTTCTCTGTGAAATCTCAAATATTGTTGTAATAGT
CTGTTTCAATCTTAAAAAGAATCA

508/615

FIGURE 504

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96889`  
><subunit 1 of 1, 339 aa, 1 stop  
><MW: 36975, pI: 7.85, NX(S/T): 1  
MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLDPIQ  
LKCVGGTAGCDSYTPKVIQCQNKGWDGYDVQWECKTDLDIAYKFGKTVVSCEGYESSEDQYVL  
RGSCGLEYNLDYTELGLQKLKESGKQHGFASFSDYYYKWSSADSCNMSGITIVVLLGIAFVV  
YKLFLSDGQYSPPPYSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQNTGHGATSGFGSAFTG  
QQGYENSGPGFWTGLGTGILGYLFGSNRAATPFSDSWYYPSYPPSYPGTWNRAYSPLHGGS  
SYSVCSNSDTKTRTASGYGGTRRR
```

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 171-190

N-glycosylation site.

amino acids 172-176

Glycosaminoglycan attachment sites.

amino acids 244-248, 259-263, 331-335

Tyrosine kinase phosphorylation site.

amino acids 98-106

N-myristoylation sites.

amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,
270-276, 278-284, 312-318

GCAAAAGGAAGGAGGGAAGCACTCCATCATCTCACTGGGAAGAACGGGCACGGGCATACCTGC
AGCTACTGGGGTTCCACTGGGCTTGAGGGTCGATTTTTTACCTTTTGAAGGACAAGATGCATT
GGAAGATGTTGCTGCTTCTGCTGTTGTATTACAATGCTGAGGCTTCTATGTGCCACAGGTGGA
GCAGGGCTGTGCTCTTCCCTGCCGCCACCGGCCAAAGAGGTCCTCATCACTGCCATTGAACC
CAGTCCTGCAGACCTCCCTGGAGGAGGTGGAGCTGCTCTACGAGTTCCTGCTGGCCGAACCTG
AGATCAGCCCTGACCTGCAGATCTCCATCAAGGACGAGGAGCTGGCCTCCTTGCGGAAGGCCT
CAGACTTCCGCACCGTCTGCAACAACGTCATCCCCAAGAGCATCCAGACATCCGCCGGCTCA
GCGCCAGCCTCTCCAGCCACCCTGGCATCCTCAAGAAAGAAGACTTTGAAAGGACAGTGCTGA
CCCTGGCCTACACAGCCTACCGCACAGCCCTGTCCACGGCCATCAGAAGGACATCTGGGCGC
AGTCCCTCGTTAGCCTCTTCCAGGCCCTGAGGCACGACTTGATGCGCTCCTCACAGCCGGGAG
TACCTCCCTGAGAGACTGGCCACACCAGGACCTCAGAGCAGGGACCAGCACAGTAATCCAGA
AAGTCTTCACTTCTACTCCATTTACAGAGACCAGCAACAAAACACTTACCGCTGACACAGAG
CAGCAGAGATCAAACAGTAACCCCGATGCTCTTTTCTCCTTGATGTTTCTGGAAGACACATC
TGATTCAATGCCATCATGTGACCTGGGCTGGAAGAAAGGGCTGGAATGGTCATTCAAGACGCCT
CCATGGGCAGAATGTTTTGCCTATGGCAGGCAGAATTCTGATATGCTTCAACCCAGAGCAGTG
GCCACACACTCAAGAGTGAGAACAGGCGTGAGCCACCGTGCCTGGCCCAGGATCTAAAACTT
TCTAAGTTTCTCCTCCATCGTTGGCATCCTCACAGCTATCTCCAATGTCACTCAAGAGACATCAA
CAGACATTTAACTGCTGCAGACTTCATTGCTCTGTACCTCACCTTGAATCTAACAAATCAAA
GTATTTCTGCAGGTCCAATGGTCTAAATCAAAATGCTTGTTAAATGACTTTTTTACAACACCCCTT
ACTTTTCTAATCCATTTCAATCTTATTTTTTTTTTATTGTGGTAAAAAACACATCACGTAAATG
TACCATCTTAACCATTTTTTAAGCATATGGTACAGCAGTGTTAACTCCATGCATGTTGTGAAAC
AGACCCCCCGAACTTTCTCATCTTGTAATTCTGAAGTTCTATACCCACCGAACAACTCCTCTT
TTCCCCCTTCCCCCTGCCTGCCCCAGCTCTTGGCACCATTATTCTGCTTTTCTGTTTTTGTGAGAGT
CTGACTACTTAAGATACCTCATACAAGCGGGATCTGGCTTACATTTCTTGAGCATTGTATTCT
GGAAAAGTGTTTTCTTCTCTGAAAAATGGGTAGAGTTCTGAAGGAGAACTACTGGTCTTATT
GTACACTTGCTGTACCTATTTTTTATTTAACAAATATTTCATCTATGGTATAATAAAGATGTCAT
GGTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

FIGURE 506

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96893
><subunit 1 of 1, 173 aa, 1 stop
><MW: 19733, pI: 8.78, NX(S/T): 0
MHWKMLLLLLLLYYNAEASMCCHRWSRAVLFPAAHRPKRSSSLPLNPVLQTSLEEVVELLYEF
LLAELEISPDLQISIKDEELASLRKASDFRTVCNNVIPKSIPDIRRLSASLSSHGILKK
EDFERTVLTLAYTAYRTALSHGHQKDIWAQSLVSLFOALRHDLMRSSQPGVPP
```

Important features of the protein:

Signal peptide:

Amino acids 1-17

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 36-40;84-88;105-109

GGCGCGGGCTGCGCGGAGCGGCGTCCCCTGCAGCCGCGGACCGAGGCAGCGGCGGCACCTGC
CGGCCGAGCAATGCCAAGTGAGTACACCTATGTGAAACTGAGAAGTGATTGCTCGAGGCCTTC
CCTGCAATGGTACACCCGAGCTCAAAGCAAG**ATG**GAGAAGGCCAGCTTGTTATTTAAAGACAT
CCTCAAATGTACATTGCTTGTGTTTGGAGTGTGGATCCTTTATATCCTCAAGTTAAATTATAC
TACTGAAGAATGTGACATGAAAAAATGCATTATGTGGACCCTGACCATGTAAAGAGAGCTCA
GAAATATGCTCAGCAAGTCTTGCAAGAAGGAATGTCGTCCCAAGTTTGCCAAGACATCAATGGC
GCTGTTATTTGAGCACAGGTATAGCGTGGACTTACTCCCTTTTGTGCAGAAGGCCCCCAAAGA
CAGTGAAGCTGAGTCCAAGTACGATCCTCCTTTTGGGTTCGGAAGTTCTCCAGTAAAGTCCA
GACCCCTCTTGAACTCTTGCCAGAGCACGACCTCCCTGAACACTTGAAAGCCAAGACCTGTCTG
GCGCTGTGTGGTTATTGGAAGCGGAGGAATACTGCACGGATTAGAAC'TGGGCCACACCCTGAA
CCAGTTTCGATGTTGTGATAAGGTTAAACAGTGCACCAGTTGAGGGATATTCAGAACATGTTGG
AAATAAAACTACTATAAGGATGACTTATCCAGAGGGCGCACCACCTGTCTGACCTTGAATATTAT
TCCAATGACTTATTTGTTGCTGTTTTATTTAAGAGTGTTGATTTCAACTGGCTTCAAGCAATG
GTAAAAAAGGAAACCCTGCCATTCTGGGTACGACTCTTCTTTTGGAAAGCAGGTGGCAGAAAAA
ATCCCACTGCAGCCAAAACATTTAGGATTTTGAATCCAGTTATCATCAAAGAGACTGCCTTT
GACATCCTTCAGTACTCAGAGCCTCAGTCAAGGTTCTGGGGCCGAGATAAGAACGTCCTCCACA
ATCGGTGTCAATTGCCGTTGTCTTAGCCACACATCTGTGCGATGAAGTCAGTTTGGCGGGTTTT
GGATATGACCTCAATCAACCCAGAACACCTTTGCACTACTTCGACAGTCAATGCATGGCTGCT
ATGAACTTTTCAGACCATGCATAATGTGACAACGGAAACCAAGTTCCCTCTTAAAGCTGGTCAA
GAGGGAGTGTTGAAAGATCTCAGTGGAGGCATTGATCGTGAATTT**TGA**ACACAGAAAACCTCA
GTTGAAAATGCAACTCTAACTCTGAGAGCTGTTTTTTGACAGCCTTCTTGATGTATTTCTCCAT
CCTGCAGATACTTTGAAGTGCAGCTCATGTTTTTAACTTTTAATTTAAAAACACAAAAAAAT
TTTAGCTCTTCCCACCTTTTTTTTTTCTATTATTTAGGTCAGTGTTTGTTTTTGCACACCAT
TTTGTAATGAACTTAAGAATTGAATTGGAAAGACTTCTCAAAGAGAATTGTATGTAACGAT
GTTGTATTGATTTTTTAAGAAAGTAATTTAATTTGTAAACCTTCTGCTCGTTTACACTGCACAT
TGAATACAGGTAACCTAATTGGAAGGAGAGGGGAGGTCACTCTTTTGTATGGTGGCCCTGAACCT
CATTCTGGTTCCCTGCTGCGCTGCTTGGTGTGACCCACGGAGGATCCACTCCCAGGATGACGT
GCTCCGTAGCTCTGCTGCTGATACTGGGTCTGCGATGCAGCGCGTGAGGCCTGGGCTGGTTG
GAGAAGGTCACAACCCTTCTCTGTTGGTCTGCCTTCTGCTGAAAGACTCGAGAACCAACCAG
GAAGCTGTCTGGAGGTCCCTGGTTCGGAGAGGGACATAGAATCTGTGACCTCTGACAACCTGTG
AAGCCACCCTGGGCTACAGAAACCACAGTCTTCCCAGCAATTATTACAATTCTTGAATTCCCT
GGGGATTTTTTACTGCCCTTTCAAAGCACTTAAGTGTTAGATCTAACGTGTTCCAGTGTCTGT
CTGAGGTGACTTAAAAAATCAGAAACAAACTTCTATTATCCAGAGTCATGGGAGAGTACACC
TTTCCAGGAATAATGTTTTTGGGAAACACTGAAATGAAATCTTCCCAGTATTATAAATTGTGTA
TTTAA

512/615

FIGURE 508

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96897
><subunit 1 of 1, 362 aa, 1 stop
><MW: 41736, pI: 8.80, NX(S/T): 3
MRRPSLLLLKDILKCTLLVFGVWILYILKLNYYTTEECMDKMKMHYVDPDHVKRAQKYAQQVL
QKECRPKFAKTSMAALLFEHRYSDLLPFVQKAPKDSEAESKYDPPFGFRKFSSKVQTLLE
LLPEHDLPEHLKAKTCRRCCVIGSGGILHGLELGHNTLNQFDVVIRLNSAPVEGYSEHVG
KTTIRMTYPEGAPLSDLEYYSNDLFVAVLFKSVDFNWLQAMVKKETLPFWVRLFFWKQVA
EKIPLQPKHFRILNPVIIKETAFDILQYSEPQSRFWGRDKNVPTIGVIAVVLATHLCDEV
SLAGFGYDLNQPRTPPLHYFDSQCMAAMNFQTMHNVTTETKFLKLKLVKEGVVKDLSSGIDR
EF

Important features of the protein:

Transmembrane domains:

Amino acids 11-27;281-297

N-glycosylation sites:

Amino acids 30-34;180-184;334-338

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 2-6;109-113;223-227

N-myristoylation sites:

Amino acids 146-152;150-156;179-185;191-197

513/615

FIGURE 509

GGGCGGACGCAGTGCAGTAAGAGCAGATGGGCGGACCCAAATTTCTTCGGCTTCACGATTTTG
CCGAGGTCTAGCCCTGCATCCAGCCTTGAAACAGGGTGGGGAGGAGGCAGAAAGGGGAGGGAC
TGCACTCCCTCTGAGCGTGCTAGCTCCGACTGCCTGACGGATCACCCCTTCCGCTCCAACATGG
CTAGTTCCCTCAACGCCGTGACTCAAGCCTGTTGTGCCAGGCAGGGCGCACTCAGCAGCGCAGC
CCACAGGTGGCGAAGGCTCCGCGAGAGGGTTCCCGCCAGGCTAGACAGTGGAGTGCCGCACA
GCGCGCCTTCCAGCCTCGCAGCCGCCACCCTAGCGGTTCCGACCCGGCGCCAGCAGGCCTGCT
TGGTTCGATCTTCGAGCCAAAGATGCGGCGAGGCTGGAAGATGGCTCTGTCTGGGGGGCTGCGG
TGCTGCCGCCGGGTACTGTCCTGGGTGCCAGTGCTCGTTATTGTCTCTCGTCTGTGCTCTGGTCC
TACTATGCCTACGTCTTTGAACTCTGCCTGGTTATTTACCTCATACTCTACCATGCCATCTTT
GTGTTCTTTACCTGGACCTACTGGAAGTCTATCTTTACACTCCACAGCAGCCAAACCAGAAG
TTCCACTTGTCTACACAGACAAGGAGCGCTATGAAAATGAAGAAAGACCTGAGGTCCAGAAG
CAGATGCTTGTTGATATGGCCAAAAAGCTACCGGTTTACACAAGAAGTGGAGGTGAGGTCAG
TTCATCCAAAGGCAGCTAGAGAGGCAGCTCAGCAAGTATCTCAGAAAGGCTAAGTCATATATG
TTCTCAAACTAGCCCTTTTTTTTTCTCCCATCTTCTGAAAACCACTATGGAGATTTTTCTCCA
CATTTTATTTCTAAAAAATTTTAAACACATATCAAAGCTGGAAGAATTGTATAGTAAACAAAC
TGTATACCCCAAACCTGGATTCTTCTGCTAACATTTTTCTGTGTTGCTATATCACATATCTATC
CACATATGCATACCTCTATTTATCTTTCTGTCAAGCCATCTTATGTTTCTGATGCATTTCAAAG
TAAATAGCTGACATCAGTAAGACATCTACCTAAATATTTTATTCTGTTTTGTAAATTTTACA
TACAAAAACATGCATAATCTTAAGGGTACCATTCCATGTATTTTGAAAAGTGTACACATCTGT
GTAATAAACCCCAATAAAATTGCCATCACCTCAG

514/615

FIGURE 510

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA98564
><subunit 1 of 1, 143 aa, 1 stop
><MW: 17255, pI: 9.99, NX(S/T): 0
MRRGWKMALSGGLRCCRRVLSWVPVLVIVLVVLWSYYAYVFELCLVIYLLILYHAIFVFFT
WTYWKSIFTLPQQPNQKFHLSYTDKERYENEERPEVQKQMLVDMAKKLPVYTRTGSGGQF
IQRQLERQLSKYLRKAKSYMFSN
```

Important features of the protein:

Transmembrane domain:

Amino acids 24-45

N-myristoylation sites:

Amino acids 11-17;12-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 511

CAGCCGGGGCG**AT**GGCGGGGCTCTGGCTGGGGGCTCGTGTGGCAGAAGCTGCTGCTGTGGGGCG
CGGCGAGTGCCCTTTCCCTGGCCGGCGCCAGTCTGGTCCTGAGCCTGCTGCAGAGGGTGGCGA
GCTACGCGCGGAAATGGCAGCAGATGCGGCCCATCCCCACGGTGGCCCGCGCCTACCCACTGGTG
GGCCACGCGCTGCTGATGAAGCCGGACGGGCGAGGTAAGGGCCGGCGCTCCTCCTGGAGCGCA
ACGGGGTCCGCAGCCCCGTTCCCAACCCTCCGATCAGCCAGGAACCCGCTGCTTGTGGCGCTGG
CCGCAGGAGAGAGGAGCCTGTCAACCCTGTGGAGAATGCACTCCCAGTTCTAGTCGTTGCCCT
TGGCACCCGCCGACACTGCTAGTGCCCCATCCCAAAGTGAGCATTTCCTTTGTGTGTAGCACA
GGATGCGGTATTTCCAAACCCCTGCCCTCGGTCTTTTCCCACTCACCGCTGCTCAGCTCTCA
AAGCCCTGCCGTTTCTCCTGCTTGGCTTGGGAAGCCT**TAG**GGAACAGAAGCTCCCTGGGAGC
ACAGAGCGGTTTTTAAACTGGCCAACACCTTAACGCCCAGAGCCGCCCTCCTCTCGCTGCCACT
TTGGAAAATAAGAGACTAGAGATTCAGTGGACGCTTCCTCCCGGCATCACAAGACTTGACTGC
TGCTTCAGTTCCTCGCTTGACCTTCATACTTTAGCCCTTTAAAGGATGTTACATAATAACAATT
AAGAGACGGCAGGGCCTTCAGGCAGACTTCTTTGGAGGGTGTCAAACGCCTTGTTTTATTAAAG
AGTGAATTTTTTAATTAAAATCATGTTTTTAAACAGAGATGGACATTTTATTGATGGAAAAAA
ATCACGTTAAGTTAGAAAGCTCTCAAAGTACCTGGTATTTACAACCTCCCTGTCAGGGAGGGC
GAACTCGATCTCAGAGTTTTATTTTCATCAGGGATTACGTTGAGGTACCCAGAAATGAGAAGA
TTTGCCCAAATGGCATATTTTAAATTTGGCCAGACCAGAACCCAGTTTCCTCTGGGATTAT
TTGTTAGTAATCGTTTTACAGGCTGAGCATTAACTAACTCCAAAGCTTGAAGGACTTTTTCTC
ATTTTCACTTGTTTTCTCTAATAAAAATAATGCTGTAATTTCAACTTCACAAGATGAGGCCTC
ATGGAAGAGTGTTTACCAAATATTAATAAATACTTTGACAGAAAAAATCAAGCGAACTCTTT
GCCAACCAAATATCATCATGACTGATGTAACAAGTAATCCAACACAGATATGAAAATCACTGG
TAAAAATCATCTCAGTTAATTCTAAAAGCAGAGCTAACCACCCCTTTTGTCTAAGGCTTTAT
GGTATTAAAAAATAAACTGTACAAAAATATAGATTTTCCCCTATCCCCTACCCCTGGAAAGTA
ATATACTGAAGTCTCATCATACTGTTTTGGGGATTCCAGTAATTTAAATCTCTAGTGAACAAA
GACCTGTTTCAAACAACCTGTGAGCTGACTGGACTATTTAAAGTAATTCTCCTTGTAAGTAC
TTTCAGAGTGAAGACAATGACGAATACTGTCTTTTACAAAGGGACTTTTTATTCCACCAACAA
ATTCTGGATTTTGGCATCAGGAAAACCACTGTTCAATTTCCAACACTATATCCAAGTTGTTTG
AGAAATTATTTAAACTCTTTAACTTAGAGGGTTTTCTTTCTCCTTTACTTGTTAAAGTGACT
ATATTACAGAGTCACTTTAAGGATTAAATTTATTGCATGCAAAGTTTCTAGATCACTGTCTAG
AAGTCAGTTAGAGTAAGTTCTTTAGTTGTCAATCAAGCATTAGTAAGGCCCTGCTTTGTGCC
CAGTGTGACCTCAACAAAGTTGGGGATATCAGAATATTCTAAGATACAGTCGTTGTTCCCAA
GAATCTTGCTTTTACATACAAGAGGTGTTGCGTTTTCATTTTGCGGCTAATGTCCAAACGCTG
GCCTCAGCCATTTACCTTGAAGATTGCAGTTGGCTTCCAACCTGGCCTCTAACTCTAATCTA
GCATTTTCCAGTCCATTGTGACAAAGTCTGCCTTCCCAGCTACTCCCAGCTGTTGGACCTGC
TGCCTTAGAACCACAGATTGGTACCTCGTGCC

516/615

FIGURE 512

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA107443
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19353, pI: 10.97, NX(S/T): 0
MAGLWLGLVWQKLLWGAASALSLAGASLVLSLLQRVASYARKWQQMRPIPTVARAYPLV
GHALLMKPDGRGKGRRSSWSATGSAAPFPSPDQPGTRCLWRWPQERGACHPVENALPVLV
VAPWHPPTLLVPHPKVSIFVVCSTGCGISKPLPSVFSHLTAAQLSKPCRFLLPWLGKP
```

Important features of the protein:

Signal peptide:

Amino acids 1-25

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 75-79

N-myristoylation sites:

Amino acids 3-9;17-23;145-151

Amidation site:

Amino acids 73-77

Leucine zipper pattern:

Amino acids 8-30

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGCGGCTGGACGAGGACGCTCAGAGCCAGCTCTCGAGAGTTCAAGCAACCGACGGTTCC
CCACTGCTCCCAGGAGCGGTTACCTGGGCACTCTGTGCCCTCCTTCTGTTCGGGCCCCA
GGCCGAGGACCTGCCAGTAGGGCTCAGTTGCCTGGAGCCCGTTAGCCCATCCCCAGTT
CACTTTGCTTGTGGGATCTCCCCGTTGCTCCTGCCCGTGGACTGAGTGGCAGGCCATCCT
ACAAGCACCCGGACACTTGACATCAGTGGTGTCAAGACAACCTCTAAGAAGGTTTTCCGTG
ATCCTGCAAGCCCTGCCTTCTTCTGGGATCCTGCCTTCAATTTGATTGCACAGGTACC
ACAGCAAGCCAGTGCTGTGTGCTCCGAGTTCCAGGGCGTCTCTCCAGCTCAGCCACTGCAC
TGAGAACATGGACTCTCTGTGGGGCCCAGGAGCCGGGAGTACCCCTTTGGGGTCCACAA
CACCCGGCTGTCCCCAGACTTGTGTCCAGGGAAGATAGTGTGTGAGGGCCCTCAAGGAGAG
CGGGGCAGGGATGCCTGAGCAGGACAAGGACCCTAGAGTCCAAGAGAATCCTGGTGATCA
GAGAAGGGTCCCCGAGGTACCCGGGGATGCACGGTCTGCATTTTCGGCCCCCTGCGGGACAA
TGGAGGCCTCTCTCCCTTTGTGCCCCGGGCCCGGGCCTCTGCAGACAGACCTCATGCCCA
GAGGTGAGAAATCAGATATAACCAGACATCCAGACCTCCTGGACGAGCTCCTGCACCAA
CCGAAATGCCATCTCCAGCTCCTACAGCTCCACGGGAGGCTTGCTGGGGCTAAAGCGGAG
GAGGGGGCCAGCCTCATCCCACTGCCAGCTGACCCTCAGTTCTCTCAAAGACAGTGAGTGA
GGACAGGCCTCAGGCTGTCTCTTCAGGTACACCCAGTGTGAAAAGGCAGCAGATATAGC
ACCAGGGCAGACACTCACCTCAGGAATGACTCCTCCACATCCGAGGCCTCTAGGCCCAG
TACACACAAGTTTTCCCTGCTGCCACGCAGGCGAGGGGAGCCTTTGATGCTGCCACCTCC
CTTAGAGCTGGGGTACCGGGTCACTGTTGAAGACCTGGACCGGGAGAAGGAGGCGGCCTT
CCAGCGCATCAACAGTGCAGTGCAGGTGAGGACAAGGCCATCTCGGACTGCAGACCCTC
ACGGCCTTCCACACTTTGTCTCACTTGCAACAGGGGCTTCTGGTCTGCCTGCCGTTTTC
TAAAGCACCCAGTATGGATGCACAGCAGGAGACACACAAGTCCCAAGACTGCCTGGGCCT
ACTGGACCCCTTAGCATCTGCTGCAGGGGTCCCCTCTACAGCTCCCATGTCTGGGAAGAA
GCACAGACCACCAGGCCCCCTGTTCTCCTCCTCAGATCCCCTTCTGCCACCTCTTCTGA
TTCCCAGGACTCAGCCCAGGTCACCTCGCTGATTCTTGCCCCCTTCCCAGCTGCAAGCAT
GGATGCGGGCATGAGAAGAACAAGGCATGGCACTTCTGCTCCTGCAGCTGCCGCAGCAGC
CCCTCCCCGCTCCACATTGAACCCACAGTTGGGGTCACTACTGGAGTGATGGAGGCCCT
TCACATTTCTGGGCCTCAGCCACAGCTGCAGCAGGTGCCAGAGGTGAGAACCAGAGATC
CCAGACCTCCTGGACCAGCTCGTGCCCCAAATGAATGCCATCTCGAGCCCCCTACAGCTC
TACGGGAGGCCTCCCGGAACAAAAGCGGAAGAGGGGCCAGCCTCATCCCACTGCCAGCTG
ACCCTCAGTTCTCTCAAACACAGTGAGTGAGGACGGACCTCAGGCTGTCTCTTCGGGTAC
ACCCAGTGTGAAAAGACGGCAGATACAGCACCAGGGCAGACACTCGCTCCAGGGGTGCG
TCCCCCAGATCCCAGGCCTCTAGGCCCCGTATATGCAAGTTTCCCCTGCTGCCACGCAGG
CGAGGGGAGCCTTTGATGCTGCCACCTCCCTTAGAGATGGGGTACCGGGTCACTGCTGAA
GACCTGGACCGGGAGAAGGAGGAGGCATTCCAGCGCATCAACAGTGCAGTGCAGGTTGAG
GACCAGGCCATCTAGGACTGCAGACCCTCACGGCCTTCCCACACTTTGTCTCACTTGCA
ACAGGGGCTTCTGGTCTGCCTGCCGTTTCTAAAGCACCCAGTATGGATGCACAGCAGGAG
ACACACAAGTCCCAAGACTGCCTGGGCCTAGTGGCCCCCCTGCATCTGCTGCACAGGCCT
GTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAACGGCATAAACCCGGGAGGCAGAGC
TTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGT
CTC

518/615

FIGURE 514

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA107786
><subunit 1 of 1, 428 aa, 1 stop
><MW: 45450, pI: 9.28, NX(S/T): 3
MDSLWGPGAGSHPFVHNTRLSPDLCPGKIVLRALKESGAGMPEQDKDPRVQENPGDQRR
VPEVTGDARSAFRPLRDNGGLSPFVPGPGPLQTDLHAQRSEIRYNQTSQTSWTSSCTNRN
AISSSYSSTGGLLGLKRRRGPASSHCQLTLSSSKTVSEDRPQAVSSGHTQCEKAADIAPG
QTLTLRNDSSSTSEASRPSTHKFLLPRRRGEPLMLPPPLELGYRVTVEDLDREKEAAFQR
INSALQVEDKAISDCRPSRPSHTLSSSLATGASGLPAVSKAPSMDAQQETHKSQDCLGLLD
PLASAAGVPSTAPMSGKKHRPPGPLEFSSSDPLPATSSDSQDSAQVTSIPAPFPAASMDA
GMRRTRHGTSAPAAAAAAPPSTLNPRTLGSLLLEWMEALHISGPQPQLQQVPRGQNQRSQT
SWTSSCPK

Important features of the protein:

N-glycosylation sites:

Amino acids 105-109;187-191

Glycosaminoglycan attachment site:

Amino acids 38-42

N-myristoylation sites:

Amino acids 15-21;130-136;180-186;307-313;361-367

Amidation site:

Amino acids 315-319

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 106-117

FIGURE 515

GTCAGGGCCAGGGTGAGCGCCCGACTCCGAGCTGTCCCCGCTCCCGGCGCGGCGCTCCGCTCT
CAGCCACCTCACGGCTGCCAGGAGTGCGCGGGAGTTTGCCCCGGAGCGCGGGGAAGTTTCCTC
CGAAGCTGCGCTCCTGGAACAGCAGCACCTGCAAGCGCCCGGCAGCGGCCCCGCGAGGTTACTT
TATGGAATTGGGCTCTTAGAGAACAGAAAAAGACTGAAGTTTTACGGGAAAAACAAATCATGTG
GTCTTCAGATTCTGAAATAAGGAGAAATGCAGCCATCTGAAATGGTCATGAACCCCCAAACAAG
TCTTCCTCTCTGTGCTGATATTTGGAGTAGCTGGGCTACTCCTCTTCATGTATTTGCAAGTCTGG
ATTGAAGAACAACATACAGGGAGAGTGGAGAAGAGAAGAGAACAAAAAGTAAC TTCAGGATGG
GGACCAGTGAAGTACTTGCGGCCTGTACCCAGAATCATGAGTACAGAAAAAATCCAGGAACAT
ATCACCAACCAGAACCCCAAGTTTCACATGCCTGAGGATGTACGAGAAAAAAGGAAAATCTT
CTACTCAATTCTGAGAGATCTACTAGGCTCTTAACAAAGACCAGTCATT CACAAGGAGGGGAT
CAAGCTTTAAGTAAGTCCACAGGGTCACCAACAGAGAAGTTGATTGAAAAACGTCAAGGAGCT
AAGACTGTTTTTAACAAGTTCAGCAACATGAATTGGCCAGTGGACATTCACCCTTTAAACAAA
AGTTTAGTCAAAGATAATAAATGGAAGAAAAC TGAGGAGACCCAAGAGAAACGAAGTCTTTC
CTTCAGGAGTTTTTGCAAGAAATACGGTGGGGTGAGTCATCATCAGTCACATCTTTTTTCATACA
GTATCCAGAATCTATGTAGAAGATAAACACAAAATCTTATATTGTGAGGTACCTAAGGCTGGC
TGTTCCAATTGGAAAAGAATTCTGATGGTACTAAATGGATTGGCTTCCTCTGCATACAACATC
TCCCACAATGCTGTCCACTACGGGAAGCATTTGAAGAAGCTAGATAGCTTTGACCTAAAAGGG
ATATATACCCGCTTAAATACTTACACCAAAGCTGTGTTTTGTTCGTGATCCCATGGAAAAGATTA
GTATCAGCCTTTTAGGGACAAATTTGAACACCCCAATAGTTATTACCATCCAGTATTCGGAAAAG
GCAATTATCAAGAAATATCGACCAAATGCCTGTGAAGAAGCATTAAATTAATGGATCTGGAGTC
AAGTTCAAAGAGTTTATCCACTACTTGCTGGATTCCCACCGTCCAGTAGGAATGGACATTCAC
TGGGAAAAGGTCAGCAAACCTCTGCTATCCGTGTTTGATCAACTATGATTTTTGTAGGGAAATTT
GAGACTTTGGAAGAAGATGCCAATTACTTTTTTACAGATGATCGGTGCTCCAAAGGAGCTGAAA
TTTCCCAACTTTAAGGATAGGCACTCTTCCGATGAAAGAACCAATGCTCAAGTCGTGAGACAG
TATTTAAAGGATCTGACTAGAACTGAGAGACAATTAATCTATGACTTTTATTACTTGACTAT
TTAATGTTTTAATTATACAACCTCCACTTTTGTAGTTTGCATTCATTTTCTAAAACCTGTATAT
ACTTAATGATGATAAGTTCAAATCAGCTGTAATTTTTCTATAATTCTCTGTATGACAGAAATT
TAACCAAGTGCAGTTGTCTTGATTTAATGTAGATTTTTTACCAAATAGTATGACACCAATTGGC
ACAAAGTTATAGGAAAATCACCTACAGGAGATGTAAACAACCTTGAGTTGCTCTAAAATGTTTG
GAAAAGAGCTGCTTTTGCATTATGAATTATATTGTTGAAGCAATAACCTAGCCAGCTGTTGCA
TTAGCTAAAGCAGCCTCTTGCAATGGTAGGAAAAAAGGATCTCAAATAGCATGAGTGTATGTC
TATATCCTGAAATTTATTGTCTAAAATGCATGAATATATTTTTTAGCAGTCTGTGGCATATTAA
TCAAACCTGTTGAATTGTTTTCTTACACCTGGAAATCTTTCTATCAACTATAATGATAAATCC
ATTTTGAAGTGATATTTTGGACTTAGGCATTTTACTTTAGATTGGAAGGCATTATGTGATTTACA
ATATGAGAATATAGCAGAAAAACCA

520/615

FIGURE 516

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108682
><subunit 1 of 1, 443 aa, 1 stop
><MW: 52021, pI: 9.63, NX(S/T): 4
MQPSEMVMNPKQVFLSVLIFGVAGLLLFMYLQVWIEEQHTGRVEKRREQKVTSGWGPVKY
LRPVPRIMSTEKIQEHITNQNPKFHMPEDVREKKENLLNLSERSTRLLTKTSHSQGGDQA
LSKSTGSPTEKLIIEKRQGAQTVFNKFSNMNWPVDIHPLNKSIVKDNKWKKTEETQEKRRS
FLQEFCKKYGGVSHHQSHLFHTVSRIYVEDKHKILYCEVPKAGCSNWKRIILMVLNGLASS
AYNISHNAVHYGKHLKKLDSFDLKGITYTRLNTYTKAVFVRDPMERLVSAFRDKFEHPNSY
YHPVFGKAIKKYRPNACEEALINGSGVKFKEFIHYLLDSHRPVGMDIHWEKVSKLCYPC
LINYDFVGKFETLEEDANYFLQMIGAPKELKFPNFKDRHSSDERTNAQVVRQYLKDLTRT
ERQLIYDFYLDYLMFNYYTTPLL

Important features of the protein:

Signal peptide:

Amino acids 1-24

N-glycosylation sites:

Amino acids 159-163;243-247;324-328;437-441

Glycosaminoglycan attachment site:

Amino acids 53-57

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 177-181

Tyrosine kinase phosphorylation site:

Amino acids 329-337

N-myristoylation sites:

Amino acids 116-122;236-242

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

521/615

FIGURE 517

GGAACCTCCAGGCACCCCTGTGTGGCCGCACTGCTCCCTCTGGCCCAACCATGCCTCTGTCCAGCCACCTGCTGC
CCGCCTTGGTCCTGTTCCTGGCAGGGTCCTCAGGCTGGGCCTGGGTCCCCAACCACTGCAGGAGCCCTGGCCAGG
CCGTGTGCAACTTCGTGTGTGACTGCAGGGACTGCTCAGATGAGGCCAGTGTGGTTACCACGGGGCCCTCGCCCA
CCCTGGGCGCCCCCTTCGCCTGTGACTTCGAGCAGGACCCCTGCGGCTGGCGGGACATTAGTACCTCAGGCTACA
GCTGGCTCCGAGACAGGGCAGGGGCCGCACTGGAGGGTCCTGGGCCTCACTCAGACCACACACTGGGCACCGACT
TGGGCTGGTACATGGCCGTGGAACCCACCGAGGGAAAGAGGCATCCACCGCAGCCCTGCGCTCGCCAACCCCTGC
GAGAGGCAGCCTCCTCTTGCAAGCTGAGGCTCTGGTACCACGCGGCCTCTGGAGATGTGGCTGAAGTGCAGGCTGG
AGCTGACCCATGGCGCAGAGACCCCTGACCCCTGTGGCAGAGCACAGGGCCCTGGGGCCCTGGCTGGCAGGAGTTGG
CAGTGACCACAGGCCGCATCCGGGGTGACTTCCGAGTGACCTTCTCTGCCACCCGAAATGCCACCCACAGGGGCG
CTGTGGCTCTAGATGACCTAGAGTTCTGGGACTGTGGTCTGCCACCCCCAGGCCAACTGTCCCCGGGACACC
ACCACTGCCAGAACAAGGTCTGCGTGGAGCCCCAGCAGCTGTGCGACGGGGAAGACAACCTGCGGGGACCTGTCTG
ATGAGAACCCACTCACCTGTGGCCGCCACATAGCCACCGACTTTGAGACAGGCCTGGGCCCATGGAACCGCTCGG
AAGGCTGGTCCCGGAACCACCGTGTGGTGGTCTTGAGCGCCCTCTGGCCACGCCGTGACCACAGCCGGAACA
GTGCACAGGGCTCCTTCTGTGTCTCGTGGCCGAGCCTGGCACCCCTGCTATACTCTCCAGCCCCGAATTCCAAG
CCTCAGGCACCTCCAAGTGTCTGCTGGTCTTCTATCAGTACCTGAGTGGGTCTGAGGCTGGCTGCCCTCCAGCTGT
TCCTGCAGACTCTGGGGCCCCGGCGCCCCCGGGCCCCCGTCTGCTGCGGAGGCGCCGAGGGGAGCTGGGGACCG
CCTGGGTCCGAGACCGTGTGACATCCAGAGCGCCTACCCCTTCCAGATCCTCCTGGCCGGGCAGACAGGCCCGG
GGGGCGTCTGGGTCTGGACGACCTCATCCTGTCTGACCACTGCAGACCAGTCTCGGAGGTGTCCACCCCTGCAGC
CGCTGCCCTCTGGGCCCCGGGCCCCAGCCCCCAGCCCTGCGGCCAGCTCGCGGCTCCAGGATTCCTGCAAGC
AGGGGCATCTTGCCCTGCGGGGACCTGTGTGTGCCCCCGGAACAACCTGTGTGACTTCGAGGAGCAGTGCAGGGG
GCGAGGACGAGCAGGCCTGTGGCACCACAGACTTTGAGTCCCCGAGGCTGGGGGCTGGGAGGACGCCAGCGTGG
GGCGGCTGCAGTGGCGGCGTGTCTCAGCCCAGGAGAGCCAGGGGTCCAGTGCAGCTGCTGCTGGGCACTTCCTGT
CTCTGCAGCGGGCTGGGGGCGAGTAGGCGCTGAGGCCCGGGTCTCACACCCCTCCTTGGCCCTTCTGGCCCCA
GCTGTGAACCTCCACCTGGCTTATTATTACAGAGCCAGCCCCGAGAGGTCTCCTGTAACCTTTGAGCGGGACACAT
GCAGCTGGTACCCAGGCCACCTCTCAGACACACACTGGCGCTGGGTGGAGAGCCGCGGCCCTGACCACGACCACA
CCACAGGCCAAGGCCACTTTGTGCTCCTGGACCCCAAGACCCCTGGCCTGGGGCCACAGTGCCCACTGTCTCT
CCAGGCCCCAGGTGCCAGCAGCACCACGGAGTGTCTCAGCTTCTGGTACCACCTCCATGGGCCCCAGATTGGGA
CTCTGCGCCTAGCCATGAGACGGGAAGGGGAGGAGACACACCTGTGGTGCAGGTGAGGCACCCAGGGCAACCGCT
GGCAGAGGCCTGGGCCACCCCTTCCACCAGCCTGGCTCCCATGCCAGTACCAGCTGCTGTTGAGGGCCCTCC
GGGACGGATACACGGCACCATGGCGCTGGACGATGTGGCCGTGCGGCCGGGCCCCCTGCTGGGCCCCTAATTACT
GCTCCTTTGAGGACTCAGACTGCGGCTTCTCCCTGGAGGCCAAGGTCTCTGGAGGCGGCAGGCCAATGCCTCGG
GCCATGCTGCCTGGGGCCCCCAACAGACCATACCACTGAGACAGCCCAAGGGCACTACATGGTGGTGACACAA
GCCCAGACGCACTACCCCGGGGCCAGACGGCCTCCTGACCTCCAAGGAGCACAGGCCCTGGGCCAGCCTGCTT
GTCTGACCTTCTGGTACCACGGGAGCCTCCGCAGCCAGGCACCCCTGCGGGTCTACCTGGAGGAGCGCGGGAGGC
ACCAGGTGCTCAGCCTCAGTGCCACGGCGGGCTTGCCCTGGCGCTGGGCAGCATGGACGTGCAGGCCGAGCGAG
CCTGGAGGGTGGTGTGTTGAGGCACTGGCCGAGGCGTGGCACACTCCTACGTGGCTCTGGATGATCTGCTCCTCC
AGGACGGGCCCCTGCCCTCAGCCAGGTTCTGTGATTTTGTAGTCTGGCCTGTGTGGCTGGAGCCACCTGGCCGGGC
CCGGCCTGGGCGGATACAGCTGGGACTGGGGCGGGGAGCCACCCCTCTCGTTACCCCAAGCCCCCTGTGGACC
ACACCCCTGGGCACAGAGGCAGGCCACTTTGCCTTCTTTGAAACTGGCGTGTGGGCCCCGGGGGCGGGCCGCT
GGCTGCGCAGCGAGCCTCTGCCGGCCACCCAGCCTCCTGCCCTCCGCTTCTGGTACCACATGGGTTTTCTGTAGC
ACTTCTACAAGGGGGAGCTGAAGGTACTGCTGCACAGTGCTCAGGGCCAGCTGGCTGTGTGGGGCGCAGGCGGGC
ATCGGCGGCACCACTGGCTGGAGGCCAGGTGGAGGTAGCCAGTGCCAAGGAGTTCCAGATCGTGTGTTGAAGCCA
CTCTGGGCGGCCAGCCAGCCCTGGGGCCATTGCCCTGGATGACGTGGAGTATCTGGCTGGGCAGCATTTGCCAGC
AGCCTGCCCCCAGCCCGGGGAACACAGCCGCACCCGGGTCTGTGCCAGCTGTGGTTGGCAGTGCCCTCCTATTGC
TCATGCTCCTGGTGTGCTGGGACTTGGGGGACGGCGCTGGCTGCAGAAGAAGGGGAGCTGCCCCCTCCAGAGCA
ACACAGAGGCCACAGCCCTGGCTTTGACAACATCCTTTTCAATGCGGATGGTGTACCCCTCCCGGCATCTGTCA
CCAGTGATCCGTAGACCACCCCAAGGCCCGCTTCCCTCACGTGACATCCAGCACTTGGTCAGACCCCTAGCC
AGGGACCGGACACCTGCCCCGCCAGGCTGGGACAGGCTGCAGGTCTCAGGATATGCTGAGGCCTGGGCGTTCCC
TGCCCTGTGCTGACTCTGTTGCTCTGTGAATAAACACCCCTGGCCCATGAGGGCCGCCCCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

522/615

FIGURE 518

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108684
><subunit 1 of 1, 1137 aa, 1 stop
><MW: 122776, pI: 6.00, NX(S/T): 4
MPLSSHLLPALVLFLAGSSGWAVPNHCRSPGQAVCNFVCDRCDSDEAQCQYHGASPTL
GAPFACDFEQDPCGWRDISTSGYSWLRDRAGAALEGPGPHSDHTLGTDLGWYMAVGTHRG
KEASTAALRSPTLREAASSCKLRLWYHAASGDVAELRVELTHGAETLTLWQSTGPGWPGW
QELAVTTGRIRGDFRVTFSATRNATHRGAVALLDDLEFWDGCLPTPQANCPGHHHCQNKV
CVEPQQLCDGEDNCGDLSDENPLTCGRHIATDFETGLGPWNRSEGWSRNRAGGPERPSW
PRRDHSRNSAQGSFLVSVAEPGTPAILSSPEFQASGTSNCSLVFYQYLSGSEAGCLQLFL
QTLGPGAPRAPVLLRRRRGELGTAWVRDRVDIQSAYPFQILLAGQTGPGGVVGLDDLILS
DHCRPVSEVSTLQPLPPGPRAPAPQPLPPSSRLQDSCKQGHLLACGDLCPVPEQLCDFEEQ
CAGGEDEQACGTTDFESPEAGGWEDASVGRLLQWRRVSAQESQGSAAAAGHFLSLQRAWG
QLGAEARVLTPLLGPSPGSCLEHLAYYLQSQPREVSCNFERDTCWYPGHLSDTHWRWVE
SRGPDHDHTTGQGHFVLLDPTDPLAWGHSALLSRPQVPAAPTECLSFYHHLHGPQIGTL
RLAMRREGEETHLWSRSGTQGNRWHEAWATLSHQPGSHAQYQLLFEGLRDGYHGTALDD
VAVRPGPCWAPNYCSFEDSDCGFSPGGQGLWRRQANASGHAAWGPPTDHTTETAQGHYMV
VDTSPDALPRGQTASLTSKEHRPLAQPACLTFWYHGSLRSPGTLRVYLEERGRHQVLSLS
AHGGLAWRLGSMQVQAERAWRVFEAVAAGVAHSYVALDDLLLDGPGCPQPGSCDFESGL
CGWSHLAGPGLGGYSWDWGGGATPSRYPQPPVDHTLGTGTEAGHFAFFETGVLPGGRAAWL
RSEPLPATPASCLRFWYHMGFPPEHFYKGLKVLHLSAQGQLAVWGAGGHRHQWLEAQVE
VASAKEFQIVFEATLGGQPALGPALDDVEYLAGQHCQQPAPSPGNTAAPGSVPAVVGSA
LLLLMLLVLLGLGGRRWLQKKGSCPFQSNTEATAPGFDNILFNADGVTLPASVTSDP

Important features of the protein:

Signal peptide:

Amino acids 1-20

Transmembrane domain:

Amino acids 1075-1092

N-glycosylation sites:

Amino acids 203-207;281-285;339-343;756-760

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 514-518;1100-1104

N-myristoylation sites:

Amino acids 32-38;55-61;61-67;106-112;116-122;336-342;350-356;409-415;
523-529;540-546;678-684;707-713;791-797;870-876;921-927;
937-943;954-960;1036-1042;1071-1077

Amidation site:

Amino acids 1093-1097

Cell attachment sequence:

Amino acids 191-194

523/615

FIGURE 519

GCAGGGGAGCTCCGAGTGTCCACAGGAAGGGAACTATCAGCTCCTGGCATCTGTAAGGATGCT
GTCCATGCTGAGGACAATGACCAGACTCTGCTTCCTGTTATTCTTCTCTGTGGCCACCAGTGG
GTGCAGTGCAGCAGCAGCCTCTTCTCTTGAGATGCTCTCGAGGGAATTTCGAAACCTGTGCCTT
CTCCTTTTCTTCCCTGCCTAGAAAGCTGCAAAGAAATCAAGGAACGCTGCCATAGTGCAGGTGA
TGGCCTGTATTTTCTCCGCACCAAGAATGGTGTGTCTACCAGACCTTCTGTGACATGACTTC
TGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCACGAGAATGACATGCGTGGGAAGTGCAC
GGTGGGTGATCGCTGGTCCAGTCAGCAGGGCAACAAAGCAGACTACCCAGAGGGGGATGGCAA
CTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCC
TGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCATGTGCCCAACAAGTCCCCCAT
GCAGCATTGGAGAAACAGCGCCCTGCTGAGGTACCGCACCAACACTGGCTTCCTCCAGAGACT
GGGACATAATCTGTTTGGCATCTACCAGAAATACCCAGTGAAATACAGATCAGGGAAATGTTG
GAATGACAATGGCCCAGCCATACCTGTGGTCTATGACTTTGGTGATGCTAAGAAGACTGCATC
TTATTACTCACCGTATGGTCAACGGGAATTTGTTGCAGGATTCGTTTCAGTTCCGGGTGTTTAA
TAACGAGAGAGCAGCCAACGCCCTTTGTGCTGGGATAAAAGTTACTGGCTGTAACACTGAGCA
TCACTGCATCGGTGGAGGAGGGTTCTTCCACAGGGCAAACCCCGTCAGTGTGGGGACTTCTC
CGCCTTTGACTGGGATGGATATGGAACACGTTAAGAGCAGCTGCAGTCGGGAGATAACGGA
GGCGGCTGTACTCTTGTCTATAGATGAGACAGAGCTCTGCGGTGTCAGGGCGAGAACCCATC
TTCCAACCCCGGCTATTTGGAGACGGAAAACTGGAATTCTAACAAGGAGGAGAGGAGACTAA
ATCACATCAATTTGCA

205 kb "GAGGAGT"

524/615

FIGURE 520

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108701
><subunit 1 of 1, 325 aa, 1 stop
><MW: 36212, pI: 8.68, NX(S/T): 1
MLSMLRTMTRLCLLFFSVATSGCSAAAASSLEMLSREFETCAFSFSSLPRSCKEIKERC
HSAGDGLYFLRTKNGVVYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQOGNKA
DYPEGDGNWANYNTFGSAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSALLR
YRTNTGFLQRLGHNLFGIYQKYPVKYRSGKCWNNDNGPAIPVVYDFGDAKKTASYYSYGQ
REFVAGFVQFRVFNNERAANALCAGIKVTGCNTEHHCIGGGGFPPQGKPRQCGDFSAPDW
DGYGTHVKSSCSREITEAAVLLFYR
```

Important features of the protein:

Signal peptide:

Amino acids 1-26

Glycosaminoglycan attachment site:

Amino acids 86-90

N-myristoylation sites:

Amino acids 23-29;88-94;127-133;136-142;265-271

Downloaded from www.nrj.com

525/615

FIGURE 521

GATCAGTGTGTGAGGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTTCTGCTCCTGCAGCTCTTCT
GTGTTGGCTGTGGATTCTGTGGGAAAGTCTTGGTGTGGCCCTGTGACATGAGCCATTGGCTTAATGTCAAGGTCA
TTCTAGAAGAGCTCATAGTGAGAGGCCATGAGGTAACAGTATTGACTCACTCAAAGCCTTCGTTAATTGACTACA
GGAAGCCTTCTGCATTGAAATTTGAGGTGGTCCATATGCCACAGGACAGAACAGAAGAAAATGAAATATTTGTTG
ACCTAGCTCTGAATGTCTTGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAAATGATTTTTTTTGTGAAA
TAAGAGGAACTTTAAAAATGATGTGTGAGAGCTTTATCTACAATCAGACGCTTATGAAGAAGCTACAGGAAACCA
ACTACGATGTAATGCTTATAGACCCTGTGATTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTG
TGCTCACACTTAGAATTTCTGTAGGAGGCAATATGGAGCGAAGCTGTGGGAACTTCCAGCTCCACTTTCCCTATG
TACCTGTGCCATGACAGGACTAACAGACAGAATGACCTTTCTGGAAAAGAGTAAAAAATTCATGCTTTTCAGTTT
TGTTCCACTTCTGGATTACGATTACGACTATCATTTTTTGGGAAGAGTTTTATAGTAAGGCATTAGGAAGGCCCA
CTACATTATGTGAGACTGTGGGAAAAGCTGAGATATGGCTAATACGAACATATTGGGATTTTGAATTTCCCTCAAC
CATACCAACCTAACTTTGAGTTTGTGGAGGATTGCACTGTAAACCTGCCAAAGCTTTGCCTAAGGAAAATGGAAA
ATTTTGTCCAGAGTTACAGGGGAAGATGGTATTGTGGTGTTTTTCTCTGGGGTCACTGTTTCAAAATGTTACAGAAG
AAAAGGCTAATATCATTGCTTCAGCCCTTGCCAGATCCCACAGAAGGTGTTATGGAGGTACAAAGGAAAAAAC
CATCCACATTAGGAGCCAATACTCGGCTGTATGATTGGATACCCAGAAATGATCTTCTTGGTTCATCCCAAAACCA
AAGCTTTTATCACTCATGGTGAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGGAGTTT
CCATATTTGGTGATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAAATAAACTTCAAAA
CTATGACAAGCGAAGATTTACTGAGGGCTTTGAGAACAGTCATTACCGATTCTCTTATAAAGAGAATGCTATGA
GATTATCAAGAATTCACCATGATCAACCTGTAAAGCCCCCTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCTATGC
GCCACAAAGGAGCCAAGCACCTGCGATCAGCTGCCCATGACCTCACCTGGTTCAGCAGTACTCTATAGATGTGA
TTGGGTTTCTGCTGACCTGTGTGGCAACTGCTATATTTCTGTTTCAAAAATGTTTTTTATTTTCCCTGTCAAAAAT
TTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTTCAAAATTCAGAAAAGACCTGATGGGGTAATCCTG
TTAATTCAGCCACATAGAATTTGGTGAAAACCTTGCTATTTTCATATTATCTATTCTGTTATTTTTATCTTAGCT
ATATAGCCTAGAATTCATGATCATGAGGTTGTGAGTATATCTCATTCTTTCGTTGTATTTTCCCTAGGTGTCTTT
ACTCTCTTCTCTCACTTTGTGACACAAGGACATGAATACATCTAAATTTTCTTATTTCTGATATGACTGTTTTGA
TGATGTCACTTCTATAACCTTAAAGTGATAGGGTGACATGCAATATGATTATTCCTGGTGTGCGCCCAACAC
ATGGATATAAAGAGGTAAAAAACTTAAATTCACAAAATTCAGTAAACCACACAAATCAGGTAAAGTGTCTATGA
GATTAGCTGGCTATGAGAAAATAATGATGTTTTCTTTTTCAATTTAAATAAGCCTTTCTACATAGCCAGCATCAG
TGATCTCAGAAAATAAATTGCTAATAATGATGACATGGCATTATGCTTAGAAAAGTTTGCTGTATTTCCATAGAC
CTCATCTAGATGTGATGGCCTACATTTCTGCCATCACTCAACCAATACTTTTTTCTGTTTTCTTGATGATAAAAA
GACCTTTCTCATGATTGCCATCAAATAACAAAAGAAAATATTTTTTTCTCACATAGAGAACATGTCAGTAAGAT
ATTCAGGTGAACAGATATTTTTTGGGATTAGTAACATTTGAAATATGTGGTGATAATTACTGAGTTTATAAAAT
TTATTTGATAGTACACTTAAAGAAGATTTATATGTTTTATTCTTTAAAAATGATGAATACTCAATTTCTTATCTC
TATAATCAAAAGTATAATTTACTGTAGAAAAATAAAGAGATGCTTGTCTGAAAGTAAGATCAGTGAACCTGCTTT
TCAGTCTCAATCTTTGAGAATTGTAAATTCATCAAATAATTGCTTACATAGTAAAAATTTAAGGTATTAGAAAAC
CTGCATAACAAATAGTATTATATATTAATATTTTGATATGTAAAGCTCTACACAAAGCTAAATATAGTGTAATA
ATGTTTACACTAGTAAGCAAATATGTTAATCTTCTCATTTTTTTACTGTGATATAATCTTAGTGATATGCCATT
AATAGTTTTAAATAAATAAATTTGGCTTATCTGGCTTTTTGAAAATTTTGAATTTCTACAGATGTGATTAGGTA
TATCTACAAATTAATTTCAATTTTAAATGATGATATAAAAAATAAATAAGTATTTTTCTTGTGTATGTATACA
ATAAATATAAATAAATTTGTTTACTGTTTTGAAAGTTTCTTAAGTTTTA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

526/615

FIGURE 522

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108720
><subunit 1 of 1, 527 aa, 1 stop
><MW: 60284, pI: 8.31, NX(S/T): 3
MRSDKSALVFLLLQLFCVGCFCGKVLVWPCDMSHWLNVKVILEELIVRGHEVTVLTHSK
PSLIDYRKPSALKFEVVHMPQDRTEENEIFVDLALNVLPGLSTWQSVIKLNDDFFVEIRGT
LKMMCESFIYNQTLMKKLQETNYDVMLIDPVIPCGDLMAELLAVPFVLTLRISVGGNMER
SCGKLPAPLSYVPVPMTGLTDRMTFLERVKNSMLSVLFHFWIQDYDYHFWEEFYSKALGR
PTTLCETVGKAEIWLIRTYWDFEFPQPYQPNFEFVGGLHCKPAKALPKEMENFVQSSGED
GIVVFSLGSLFQNVTEEKANIIASALAQIPQKVLWRYKGKKPSTLGANTRLYDWIPQNDL
LGHPKTKAFITHGGMNGIYEAIYHGVPMVGVPIFGDQLDNIAHMKAKGAAVEINFKTMTS
EDLLRALRTVITDSSYKENAMRLSRIHHDQPVKPLDRAVFWIEFVMRHKGAKHLRSAAHD
LTFWQHYSIDVIGFLLTCVATAIFLFTKCFLFSCQKFNKTRKIEKRE

Important features of the protein:

Signal peptide:

Amino acids 1-21

Transmembrane domain:

Amino acids 489-510

N-glycosylation sites:

Amino acids 131-135;313-317;518-522

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 67-71;340-344

Tyrosine kinase phosphorylation sites:

Amino acids 122-131;136-144

N-myristoylation sites:

Amino acids 19-25;276-282;373-379;377-383

Amidation site:

Amino acids 338-342

FIGURE 523

GGCTGCGGGGTCGGCACGGAAGATGCACGCGAGGCTCCTGGGGCTCTCGGCCCTGCTGCAGGC
 GGCCGAACAGAGCGCGCGCCTTTACACCGTGGCTTACTACTTCACCACAGGACGGCTTCTGTG
 GGGGTGGCTGGCCCTTGCTGTCTCCTGCCCGGGTCTTGGTCCAGGCCCTGAGCTACCTGTG
 GTTCCGAGCAGACGGGCATCCAGGGCATTGCTCCTTGGTGATGCTGCACCTCCTACAGCTTGG
 TGTTTGAAGCGGCACTGGGACGCTGCACTGACCAGTCTGCAGAAGGAACTGGAGGCTCCCCA
 CCGAGGCTGGCTGCAGCTGCAGGAGGCCGACCTGTCGGGCCCTTCGACTCTTGAGGCCCCTGCT
 GCAGACTGGGCCCCACCTGCTGCTTCAGACATATGTTTTTCTAGCCTCAGACTTTCACAGATATT
 GTGCCAGGGGTGAGCACCTGTTTTTCTGGTCTCCTCACTCTCCTGGGCACTGGTGTCTTACACT
 CGCTTCATGGGCTTCATGAAGCCAGGCCACCTGGCCATGCCATGGGCCCGCCCTCTTCTGCCAG
 CAGCTCTGGAGGATGGGCATGTTGGGAACCCGCGTGCTGAGTCTGGTTCTGTTCTACAAAGCC
 TACCACTTTTGGGTTTTTGTGGTTGCAGGTGCCCACTGGCTGGTGATGACATTCTGGCTTGTG
 GCCCAGCAGAGTGACATCATCGACAGCACCTGCCACTGGAGGCTGTTCAACCTGCTCGTGGGG
 GCCGTGTACATCCTCTGCTACCTCAGCTTCTGGGACAGCCCTTCTAGAAATAGGATGGTCACG
 TTCTACATGGTCATGCTGTTGGAGAACATCATCCTGTTGCTGTTGGCCACCGACTTCTCCAG
 GGGGCATCGTGGACCAGCCTGCAGACCATAGCTGGGGTCTGTCTGGATTCTGATTGGCAGT
 GTCTCACTGGTAATTTATTACAGCCTGCTGCATCCAAAATCCACAGACATCTGGCAGGGCTGC
 CTAAGGAAGTCTGTGGCATTTGCAGGAGGTGATAAAACAGAGAGAAGAGATTCTCCCCGGGCC
 ACAGATCTAGCTGGGAAGAGAACCAGAGAGCTCAGGCTCATGCCAAGGGGCAAGTTATGAACCA
 ACCATTTTAGGGAAGCCCCCTACCCCTGAGCAGGTCCCCCAGAGGCTGGGCTGGGGACCCAG
 GTTGCTGTGGAGGACTCTTTCCTCAGTCATCACTACTGGCTGTGGGTGAAACTTGCCCTAAAA
 ACAGGAAATGTGTCTAAGATCAATGCCGCCTTTGGAGATAACAGTCCTGCCTATTGTCCACCT
 GCATGGGGGTGAGTCAACAGGACTACCTGCAGAGAAAGGCCTTGTCTGCCAGCAGTACAGAGCTC
 CCATCCTCATCCCGTGACCCCTCAACCTTAGAGAACAGCTCTGCGTTTGAAGGTGTCCCTAAA
 GCAGAGGCCGACCCATTGGAAACCTCAAGTTACGTATCTTTTGCCAGCGATCAGCAGGATGAA
 GCACCTACCCAGAACCAGCAGCCACGCAGGGGGAGGGCACCCCAAAGGAAGGAGCTGACGCT
 GTTTCTGGGACACAGGGGAAGGGGACAGGTGGGCAGCAGAGAGGAGGGGAAGGACAGCAGAGT
 TCCACGTTGTACTTCAGCGCCACTGCAGAAGTGGCCACATCCTCACAACAAGAAGGCAGCCCA
 GCTACTCTGCAAACGGCCCACTCTGGAAGGAGGCTGGGAAAGAGCAGCCCTGCCAGCCTGCA
 TCGCCCCACCCAGTGGGCTTGGCGCCCTTCCCCGACACCATGGCCGACATTAGCCCCATCCTA
 GGCACAGGCCCATGTAGAGGCTTCTGCCCCAGTGCAGGCTTCCCTGGAAGAACCCTCAGTATC
 TCAGAGCTAGAGGAGCCGCTGGAGCCCAAAGGGAGCTAAGTCACCATGCAGCTGTTGGTGTG
 TGGGTGTCAATTGCCACAGCTGAGGACTGCCCATGAGCCCTGCCTCACGTCCACCCCTAAGTCT
 GAGTCTATCCAAACGGACTGCAGCTGCAGGGAACAGATGAAGCAAGAGCCGAGTTTTTTTCATC
TGACCACAGTCATGGTGGGATAAGACAACAGGCTGACAAACCAAGCTGGCCATTTGGTACCGT
 GAGAAAGGAAATCCCACTTCTGACACCTGTGTCTTGGGCACATCACTGTACCTCTGAATCT
 CCATCTGCATCCCTGAAAAATGAAGAAACAGGGCTGGATGATTTTGCAGGTCCAATGCAACA
 TCACAGACCCACCCATGCATAGGAGAGACTCTAACATACTTTAGAGGAGGAGAAAGAGATTCT
 CAGTCAAAATTGTCTGCTACCTTTTATGAGCTGTAGGTTCCCTTATTTTATCTTTTTGCTGTG
 GCTTCTAGGAAACACAAAGGTAAAACCCAGATTCCATATTTTATTTGAGGTTCTTGTTACAATT
 AGCTTTGCCTCACATTTTAGCGGTTATGAATCTCATTTTAAATATATTCTAACTGTATTATGTTA
 TGAAATCTCTTGGTAAGATAATTTGCATGCTTTCTGGGAGTAGGTAAGGCCTGTGTGCTTGTA
 ATAATAACATAACTGAAAGTGCAAATGTCA

528/615

FIGURE 524

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108726
><subunit 1 of 1, 686 aa, 1 stop
><MW: 74981, pI: 6.60, NX(S/T): 2
MHARLLGLSALLQAAEQSARLYTVAYYFTTGRLWGLALAVLLPGFLVQALSYLWFRAD
GHPGHCSLVMLHLLQLGVWKRHWDAALTSLQKELEAPHRGWLQLQEADLSALRLLEALLQ
TGPHLLLQTYVFLASDFTDIVPGVSTLFSWSSLSWALVSYTRFMGFMKPGHLAMPWAALF
CQQLWRMGMLGTRVLSLVLFYKAYHFWVFVAGAHWLVMTFWLVAQQSDIIDSTCHWRLF
NLLVGAVYILCYLSFWDSPSRNRMVTFYMVMLENIILLLLATDFLQGASWTSLQTIAGV
LSGFLIGSVSLVIYYSLHHPKSTDIWQGCLRKSCGIAGGDKTERRDSPRATDLAGKRTE
SGSCQGASYEPTILGKPPTPEQVPPEAGLGTQVAVEDSFLSHHHWLWVKLALKTGNVSKI
NAAFQDNTPAYCPPAWGLSQQDYLQKALSAQQELPSSSRDPSTLENSSAFEGVPKAEAD
PLETSSYVSFASDQQDEAPTQNPAATQGEQTPKEGADAVSGTQGKGTGGQQRGGEGQQSS
TLYFSATAEVATSSQQEGSPATLQTAHSGRRLGKSSPAQPASPHVGLAPFPDTMADISP
ILGTGPCRGFCPSAGFPGRTLSISELEEPLEPKRELSHHAAVGVVWVSLPQLRTAHEPCLT
STPKSESIQTDCSCREQMKQEPSFFI

Important features of the protein:

Signal peptide:

Amino acids 1-17

Transmembrane domains:

Amino acids 35-50;269-287;293-313

N-glycosylation sites:

Amino acids 416-420;467-471

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 344-348

N-myristoylation sites:

Amino acids 188-194;288-294;299-305;335-341;338-344;362-368;
390-396;473-479;529-535;536-542;558-564;603-609;
643-649

Amidation sites:

Amino acids 354-358;568-572

Leucine zipper pattern:

Amino acids 112-134

FIGURE 525A

AGTGCCCTCTGTCAATTTAATCTCCACGGAGTCCTGAAGGTGCTTCCCAGGTTTGTCCCCATCACACAGATGAGGC
 AATCGTTTCTGTTAAGACTGTCCTTGGGTCAAT**ATGT**CAGCAATGAAATCGGTGCTGCCATTACTAAACCTTACT
 GTGTGCTGGCATTGTGTATGCATGCATGTGTGTGCGTGACACGTGTGTGTTTGCCTGTACATGTGCATGTGTGTG
 TTGTGTGCCTGTGTGTGCACATGTAGAAAGAAAGTGATGTGTGGGAATGGAGAAT'TCCAACCCAGGAGGAGACTG
 TGCCTGGGGCTGCCACGAGAAGTGGTGACACTTCGAGAAACGGGTTCCAAATGCACGTTGCCCTTCCAGTTCTCTG
 TGTGACCTTGGGCAAGTCACTTCAGCTCCC**TGAG**CTTTGTTTTTAAAAATATTTTTTAAATGTATAAAACCATGG
 ACCATTACATATGAAGAGAAATGTGTGTGCAAACATTCAGTTAATAATCACAAAGGTGGAGGAGTGCCTGCTCAGA
 CCCAGAGCTGACACAGGGAACCTCTCCAGAAGGCTGCAGGGCTTCCCTCCCAACCTCCAAACGGCCCACTTGCTG
 AGCCCGTGCTTATCTGTTTCAGTGGAAATGCCATGTGCCACGTTCCCTCAGAAGTAAAGCCACTAGGTAAGTGTGAC
 ACTCTGTGCCAGCTTGGTGCAAAGCCCCCTTGTGTGTGCTATCTTATTGACCTCCCAAATAGCATGGTAAGGTCA
 GTACTGTCACTTCCCCCACTTTGAAGATGAGGAGAGCACAAATCTAGATGGAAATGGAGGTACGCAGTGGAAA
 TAGGATCCAGACAGATTAATCCAATCTCAAGCCTGAATTCTTCCATTCCACGCTACGCTTGAAGCTCAATCTCTC
 TTCCTGGTTGATTCTCCCCACTTCCCCACCCCCAGATATATCCCATCGCTGCTTGGTGGACAGTAGCCATGACTG
 GGTTTTGGTAAAGGTTGCTGAATAATCAGGCTGCTGGTTAGTTTTTACATTTACCTTTCCAGTGAATGGGGC
 CCCATGAAAAGGCAGCTCAAGTTGTAAATTACTCAAAGGAAGGACAGAAAGGTCTTCTGTTTGCACCTACCCTA
 AGGATTTGGGGTAGACACTGGGAATTTACTAATTATGAATTCAGTGCTTTCCCTTGCTGAAAGAGAGGCGTGGAA
 TCAACGCTGAGTGAAGGCATCAAGTTTAAAGCTGCTAATTACTTCTGATCATGCAGAATAAAAGCTACGTCCCTT
 GAAATACACCAGGCAGCTAAACATAATCTTTGCGTTTCCGTAGTGTTGGTTAAGGAATCCAGATGTTACTGCAAT
 AACCACTCCATAAACAAAAGGAACACCCAGCTGTGAGAACTGGCTTCTCAGCATTTCGTCCCAGCAGAGGCTCTTC
 CGGGGCCAGCCCTGGAAGAACCCATCAGGGTTCTGATGGTTGCCCTGTTTCAGCACAGCCCTTATTTGGCAGGCAG
 ACGGCTACGGGCACAGCCACAGGCTGAAGGTGAGTCCAGCACACAACTTTCTGACAGTGAACAGGAGTAAACATG
 GGACCCACCCGAAACCTTTGTCTGTTGACTTCTTAGCAAATGGAGGCAGCTCTAGGCTCTGGAGAGTTCGGGTAT
 AGGAGACCATGACTTGAGCAGACTGATATAAGTGAATGCAAACATATTTAGATGGCACAACCTTAATTTAGATTT
 ATCAGTGCTAATATAGAAAAAGCTAGTATTTATTGGGGCTTATTAGATTTTTTAGTCTGAATCCTCACAACCTTACG
 AGGGGGTTCGTTTTTACAGACTATGATCTTGCATGATTTCCCCAAAGATGCTCATTAAGTATATGGTGAAAGTAGA
 ATTTGAATACAGAAGACCTGGTTCTGCTACTTTCTGTGTTTCTATTTGGTTCAAACCAGCCCTTTCTCTTTCAA
 ACAACTTCAGTGCAATTCATGGTTTTTGGAAAAATAAATTTGATTTTGAGATTCAGACAATAAGTGCATTTTAAATG
 TTTATTCTTTTATCTTGAAAAACTGATATATTTATGAAATGATATGTGCTCACTCAGTGTCAACACTTCAAACAA
 CACAGACAGTACAATGACAAATTGGAGATCAGCTCTAATCTCGGCCCCCAATTTAATGCATTGCTGAATATTTCTT
 CTGAACATAGTCCATCCCACACTGTCCCATGACACAAGACGCTCCAAGGGGCTGAAGATAGAGGGACTTCTGCAG
 TCAAGAGAGCTGGGAACTCTTGGACAGTCACAATGTGCATTTGGGTATTAAGGCTCTGCAAAGTTCTGCACCA
 AATAAACCCCTTGGATTGGCTTGATCCAATGCCATGTTTCCAAACCTACTTGCCCGTGGGACACCTTAGTCCATA
 ACACAGGTTGGCATTTCCTCTAGAGAGTGTGCTGTGAAAAACACTGGTCTCACAGCACCGTGCATTTCAATCCAGCA
 GGTATTTACCAAGCAGGGACTTTGGGCCAGGTCCGTGCTAGGCTCTGCAGGTGGACCAGCCAGCCCTGACCTCCA
 TGGTGTCTCTTCTCATGGGAGAGGCTGCACAGCAGTCATTGAGAAAACGAAGAAACACACAGGTACTTTTCAATG
 CTGATAATGACTACCATGTGCTAAAAGAGCTCCAGGTGTTGTCTGTTTTGAGACAATCTTCTCGACAATGAGATA
 GAATGAACCATGCAAACCTTTGGGGCTACGATGGTTTTAGGAAAGAGCTAGAGTGAAAAATCCTTTGACATATAT
 ACATACAAATAAAGATACATGTGTATAATTTTATGTAATTGACATCACATTATATGCTGGCACCTGCCCTTTTTTG
 CTTAATGAAAATGTCAAAGGTAATATTACCAATCAATCAATATATTTTACCATCGTGTTTAAATTTCTGCAGAGCA
 CTTATTCAAGCACATACTGCATTTTTCTTAGTCACTCACTGGTGGGAATAGAAAGTTTTCTCAGGCATGCATTTT
 TCCCCTTCTGAAATGATCTCCTTAGAATAAATTTCCCAAATTTGATTTGCTGGGTCAAAGGCATGAACATTTTCA
 CATTTTTATACAATAATGTCAAACCTACCTTCCGGAATGATGCTCACTTTACTTTCCCTCCAAGACTGTGTGAAA
 ATGCCCATTTTCTAAATGCTTACTATGACTGGTTTCCAACCTACATTTTAAATCTTGTTCATCTGATAGGCAAA
 AAATGATATTTAATTTTTATTTGATTGATAATGACCTTGAACGTGCCCATTAGCCTTTTGCATGTATTCTTTTAT
 GAAACATCTGTTCCATCCTTCGTCAATGTTCCCTCTTCATATGTTAACTTTTCTTATTGATTTGTTAGAGCACTT
 TGTATATTGTGAATATTAGCTTTTGTGATCAGAATTATGGAATTGTTGTTGTCTTTAGTTTTGCTTATCAAGTTT
 AAAGCCATTTCAGAGATGTTGTAAATGTGTATGTTGTTAAATTTATGTCTTATTTTGTGTTTTCTGATGCTCATAT
 GTTTAGAAATGGTCAAGCAATCCCACTTATGATTACATAAATATTCTCCCATATTTGCTCTTAGCATTTTTTTTCT
 TTTTCATGTAACCCCTTTGTTCTTTTCAAGAAATTTACTTTGCGGTAAGAAATGGGCTGGCTTCCAGTTTATGTTTC
 CCAAATGGTGATTGAGTGTTCATTTCCATATCTCCCTTATTAGAAATGACCACTATATTATGTTCTAAATA
 TCTCGCTACTGTGCTCCCTTCCATATCTCAGTTACTCTCTTTGAGCTATCTTTTGATTCTTTTCAACAC
 ACTGCTTTACTGAACTGTCATCATCTTATACATTTTAAATACTCAGCAAGACAAGTTTCTCAATGCCACTCTTTT
 TCAGAGTTTTTCTGGTGGTTGTAAGATGTTTATCTTCTGGATAAACTTTAGAAATCACTCTTTTTTGTCCAAGGTA
 AAATATATCCACATTGAGATCATACTGAATATACAGACTAATTCAGGAAAAATGTATGTCTTTATTGCATTGA
 GTCTTCTTATCCAATAAAAAAGATATGAATTTCCATGTATTGAAATCTTCACTGAGACTTATTTTTGGCTTTTCA

530/615

FIGURE 525B

CATGTCCTGCAAATGTATTGTTAAATTTATTTTTAGGTATTTTAGGGGAAATGATTTTCTAAAGTTTGTATTTTC
TAGCTTGTTATAATTTACATATGAGATAGTCATTGTTGTATATTATTTATAACTGATCATATTACTGTATTTGTA
TTGTTTTAATAGTTTTTCTATTATTTTGGGTTTTCTGGAATACAACCTTATTATCTACAAATTATGATTGTTTT
GCCTTTTCCAATGTTTCATAACTGTTTTTATATTCTTGTCGATTGCTTTGTTTCAGCACTTCTAGAATAAAGTCAT
GCAATACTAATGA

205.25.0.150

531/615

FIGURE 526

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108728
<subunit 1 of 1, 100 aa, 1 stop
<MW: 10922, pI: 8.81, NX(S/T): 0
MSAMKSVLP LLNPYCVLAFVYACMCVRAHVCVCVYMCMCVLCACVCTCRKKVMCGNGEFQ
PRRRLCLGLPREVVTLRETGSKCTLPSSSLCDLGQVTSAP

Important features of the protein:

Signal peptide:

Amino acids 1-28

N-myristoylation sites:

Amino acids 80-86;94-100

Prokaryotic membrane lipoprotein lipid attachment sites:

Amino acids 13-24;34-45

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 527

GTGAGACTTCTTTCTTCATTGTGGCTAGCTTTGAAAAGACCCTCTGAACTTCCTAAAGATATC
AAGATGATATCACCAGACTTGCCCTTTTTGACAATTGTCTTGATCATAGTTAGTTGGACAAC
TGTGGAGCACTAGCCATACTTCTTTCTTATCTTTACTATGTGTTTAAGGTTGTTTCATCTGCAA
GCCAGCTTAACAACTTTTAAGAATAGCCAGCCTGTGAATCCCAAACACTCTAGAAGAAGTGAA
AAGAAATCCAATCATCATAAAGACTCCTCAATACACCATCTTCGTTTATCTGCCAACGATGCT
GAAGATAGCCTTCGCATGCACAGTACTGTGATTAACTTACTAACATGGATTGTATTACTCAGC
ATGCCTTCTCTAATTTATTGGCTAAAGAATCTTAGGTATTATTTTAACTTAATCCTGATCCA
TGTAACCTTTGGCATTATCCTTATTCGACTATGGCAATTCTTGGAATACTTACACTGTT
TCAATAAAATCAAGTAAATTGTTGAAGACTACTTCACAATTTCCACTTCCTCTGGCTGTTGGT
GTGATTGCTTTTGGGTCAGCACATTTATATAGGCTTCCATGCTTTGTCTTCATTCCCTCTTTTA
CTCCATGCATTATGCAACTTTATGTAAGATTGGACTTAAGGAATGATGAAGATAATTTATGTG
TTTAGGGCCAGTGATAAGAGGGAACACACAGATCCATCAGTATGGACAGCAAGATCCTTTGGA
GAAGACAAGTCTATTTTTACAATATTGAAAATAGGAAATTAGTTTTGTAATGTTTGAGGGAAG
TAGTTGAAGCATGGTTTTGTTTTGTGGTGTGGAATCCATGTACTAATCATTTTTGAAAAATTC
ATGAAGGGATATATGGTGATCACTATCATTGAGGACTCCTGTGCATATAAAATAGTCTGTTTT
ATCAACTGTAAA

533/615

FIGURE 528

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108738
><subunit 1 of 1, 196 aa, 1 stop
><MW: 22225, pI: 9.90, NX(S/T): 0
MISPDLPFLTIVLIIVSWTTCGALAILLSYLYYVFKVVHLQASLTTFKNSQPVNPKHSRR
SEKKSNNHHKDSSIHHLRLSANDAEDSLRMHSTVINLLTWIVLLSMPSLIYWLKNLRYYFK
LNPDPCCKPLAFILIPMAILGNTYTVSIKSSKLLKTTSQFPLPLAVGVIAFGSAHLYRLP
CFVFIPLLLHALCNFM

Important features of the protein:

Signal peptide:

Amino acids 1-25

Transmembrane domains:

Amino acids 91-108;128-143;167-186

N-myristoylation site:

Amino acids 141-147

2025-10-10 10:00:00

CGAGACGCCGGGTCCCAACC**ATG**GCCCGCGAATTATTCCAGTACCAGTACCCCGGAGAACATGTCA
AAGTTAAAACCAGCTCCCAGCCAGGCTTCCTGGAACGGCTGAGCGAGACCTCGGGTGGGATGT
TTGTGGGGCTCATGGCCTTCCTGCTCTCCTTCTACCTAATTTTACCAATGAGGGCCGCGCAT
TGAAGACGGCAACCTCATTGGCTGAGGGGCTCTCGCTTGTGGTGTCTCCCGACAGCATCCACA
GTGTGGCTCCGGAGAATGAAGGAAGGCTGGTGCACATCATTGGCGCCTTACGGACATCCAAGCTT
TTGTCTGATCCAAACTATGGGGTCCATCTTCCGGCTGTGAAACTGCGGAGGCACGTGGAGATG
TACCAATGGGTAGAAACTGAGGAGTCCAGGGAGTACACCGAGGATGGGCAGGTGAAGAAGGAG
ACGAGGTATTCTTACAACACTGAATGGAGGTGAGAAATCATCAACAGCAAAAACTTCGACCGA
GAGATTGGCCACAAAAACCCAGTGCCATGGCAGTGGAGTCATTTCATGGCAACAGCCCCCTTT
GTCCAAATTGGCAGGTTTTTTCTCTCGTCAGGCCTCATCGACAAAGTCGACAACTTCAAGTCC
CTGAGCCTATCCAAGCTGGAGGACCCTCATGTGGACATCATTGCGCGTGGAGACTTTTTCTAC
CACAGCGAAAAATCCCAAGTATCCAGAGGTGGGAGACTTGCGTGTCTCCTTTTTCTATGCTGGA
CTGAGCGGCGATGACCCTGACCTGGGCCCAGCTCACGTGGTCACTGTGATTGCCCGGCAGCGG
GGTGACCAGCTAGTCCCATCTCCACCAAGTCTGGGGATACCTTACTGCTCCTGCACCACGGG
GACTTCTCAGCAGAGGAGGTGTTTTCATAGAGAACTAAGGAGCAACTCCATGAAGACCTGGGGC
CTGCGGGCAGCTGGCTGGATGGCCATGTTTCATGGGCCTCAACCTTATGACACGGATCCTCTAC
ACCTTGGTGGACTGGTTTTCTGTGTTTTCCGAGACCTGGTCAACATTGGCCTGAAAGCCTTTGCC
TTCTGTGTGGCCACCTCGCTGACCCTGCTGACCCTGGCGGCTGGCTGGCTCTTCTACCGACCC
CTGTGGGGCCCTCCTCATTGCGCGCCTGGCCCTTGTGCCCATCCTTGTTGCTCGGACACGGGTG
CCAGCCAAAAAGTTGGAGTGA**AA**AGACCCTGGCACCCGCCGACACCTGCGTGAGCCCTGAGG
CTGGTTGTACAATGCCACGCCTGCCTGGCTGCTTTCACCTGGGAGTGCTTTTCGATGTGGGCA
CCTGGGCTTCTTAGGGCTGCTTCTGAGTGGTTCTTTCACGTGTTGTGTCCATAGCTTTAGTCT
TCCTAAATAAGATCCACCCACACCTAAGTCACAGAATTTCTAAGTTCCCAACTACTCTCACA
CCCTTTTTAAAGATAAAGTATGTTGTAACCAGGACGTCTTAAATGATTCTTTGTGTACCTTTTC
TGTCATATTTCAGAAACCGTTCTGTGCCTGCTGGGAGTAATTCCTTTAGCAATTAAGTATTTGG
TAGCTGAATAAGGGGTGAGAACTTCTGAAACCAGAGATCTGTAATCATCTCTATTGGCCTGGG
GTGCCCTGTGCTATAAATGAGTTTCTTCACATGAAAAACACAGCCAGCCCAAGATGACTTATCT
GGGTTTAGGATTCAATAGTATTCACCTAAGTCTTATTACATGAGCAATTTTCATCAAATCTCCA
AACTCTTAAAGGATGCTTTTCGAAAAACACGCTGTATACCTAGATGATGACTAAATGCAAAATC
CTTGGGCTTTGGTTTTTTTTCTAGTAAGGATTTTAAATAACTGCCGACTTCAAAAGTGTTCTTA
AAACGAAAGATAATGTTAAGAAAAATTTGAAAGCTTTGGAAAAACCAATTTGTAATATCATTG
TATTTTTTTATTAAAAAGTTTTGTAATAAAATTTCTAAATTATCA

535/615

FIGURE 530

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108743
><subunit 1 of 1, 400 aa, 1 stop
><MW: 44876, pI: 8.32, NX(S/T): 2
MAANYSSSTSTRREHVKVKTSSQPGFLERLSETSGGMFVGLMAFLLSFYLIIFTNEGRALKT
ATSLAEGLSLVVSPDSIHVAPENEGRLVHIIGALRTSKLLSDPNYGVHLPVAVKLRRHVE
MYQWVETEEESREYTEDGQVKKETRYSYNTEWRSEIINSKNFDREIGHKNPSAMAVESFMA
TAPFVQIGRFFLSSGLIDKVDNFKSLSLSKLEDPHVDIIRRGDDFFYHSENPKYPEVGDRL
VSFSYAGLSGDDPDLGPAHVVTVIARQRGDQLVPFSTKSGDTLLLLLHHGDFSAAEEVFHRE
LRSNSMKTWGLRAAGWMAMFMGLNLMTRILYTLVDWFPVFRDLVNIGLKAFACVATSLT
LLTVAAGWLFYRPLWALLIAGLALVPILVARTRVPAKKLE

Important features of the protein:

Transmembrane domains:

Amino acids 34-53;365-388

N-glycosylation site:

Amino acids 4-8

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 140-144

Tyrosine kinase phosphorylation sites:

Amino acids 99-107;220-227

N-myristoylation sites:

Amino acids 35-41;93-99;310-316

Cell attachment sequences:

Amino acids 221-224;268-271

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

536/615

FIGURE 531

AAAAAAAAAAAAAAAAAGAAGCTCTTATGCCAGGAACCTGGAATGGAGACCAAATATATATTG
GTTATATCATAGTATCACAGGGTTACTTTGGCATTGTTGGGAACTTGAGAGAA**AT**GGGCAATAA
CTGTTACTTTAAAAGCTTGGGTGCTGTGATTCTGCCTTCAGCCTCAGCCACTTTTGTGGTGCT
TTGCGTGGCATCAGTACCTCCACTGATTCTTCTGTCTTTCCCTCTCTCTCTTCCCCCCTCTTT
CCCTTCTGTTTTTCTCAGATCTAAGGGTTATAATGGAGGGGCAAACCTGCCTGGCTATTTCAGA
TAAGACTTCACTGAGTGACTGTTTCAAGCCCATGATTTACCCTGCAGTTTAAACAGGCTCAGGAAT
TAGGTCGCATCAGTTGAGCGCGGGTCACTTAGGCCTA**TAA**TCATCATCAGACGGCAATTAAAG
GACCATTTCTGCCTTTTTCACTATTACATCCCCGCCTGTAGCCAGCCTGCCATACAGTAGA
TACTCAATAAATATTTGCTGAATGATAACCAATAA

20570 " 6666001

537/615

FIGURE 532

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108758
<subunit 1 of 1, 100 aa, 1 stop
<MW: 10316, pI: 8.52, NX(S/T): 0
MGNNCYFKSLGAVILPSASATFVVLCVASVPPLILLSFLSLFPPSFPSVFLRSKGYNGGA
NCLAISDKTSLSDCSAHDLPCLSLTGSGIRSHQLSAGHLGL
```

Important features of the protein:

Signal peptide:

Amino acids 1-47

N-myristoylation site:

Amino acids 58-64

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 533

CGGGGTGTACGAAAGAGAAACCCGGAGGGCGCCGGGGACTGGGCGGGGTCTGCAGGGCTCAG
 CTGAGCCCATGAGCTCCAGAGCTAACCCCTGAACACCCAGGCGGGCAAAGGGCTGATGTGCG
 TAGTCCCCATCCTGGAGGGGCAGGCTCTGCGCATCTGCTCCTGGCATGGCGCTGCGGCACCTC
 GCCCTCCTGGCTGGCCTTCTCGTGGGAGTCGCCAGCAAGTCCATGGAGAACACGGCCCAGCTG
 CCCGAGTGCTGTGTGGATGTGGTGGGCGTCAACGCCAGCTGCCAGGCGCAAGTCTGTGTGGT
 CCAGGCTGTTACAGGCGCTGGAACGCGGACGGGAGCGCCAGCTGCGTCCGCTGTGGGAACGGA
 ACCCTCCCAGCCTACAACGGCTCCGAGTGTAGAAGCTTTGCTGGCCCGGGTGCGCCATTCCCC
 ATGAACAGAAGCTCAGGGACCCCCGGGCGGCCACATCCTGGGGCTCCGCGCGTGGCCGCCTCC
 CTCTTCCTGGGCACGTTCTTCATTAGCTCCGGCCTCATCCTCTCCGTAGCTGGGTTCTTCTAC
 CTCAAGCGCTCCAGTAAACTCCCCAGGGCCTGCTACAGAAGAAACAAAGCTCCGGCCCTGCAG
 CCTGGCGAAGCCGCTGCAATGATCCCCCGCCACAGTCCTCAGACGTGGGGTCTGCAGGAAAG
 GAGGACCCACCACGACAGGGCAGACCCCCAATACCTGCTCCTCCTTGAAGTCCAGCTCCACCC
 GAGGACAGACGCAGCCGGCCTCCGCCAGGGCCCTCCTGAGCAGCCATCGCTTCAGTGGTGTGCTGG
 GTCAGGCGGACCCAAGAGTCAGCCCGTACGGAAGCCGCGCTACGTACAGGCGGGAGCGGCCCT
 GGACAGGGCCACGATCCCGCTGCCTTCCCGGGGGAGGCCCCTATCAGCAATGTCTGACCTGG
 AGGCCGAGACCACGCCACGCACTTGCGGGCAGGGACCCGGAGGCGGACCCCTTGCGGGGAACC
 AGCACAAAGTGTTGGCATCGCCCGCGCCCGGGACAGTCCTGGGCACAGCCTCGGCTCTGGGT
 CCTCCGCCTCCAGCGACGGACGCCAAAGGGTCCCGGGCCGCCTGAGGCTCCTCCCCACCAC
 AGCCATCTCGTTTATCGGACCAGGAGCAGGCATCCATGAGACCTCAGAGCTTCAGATCGAGGC
 CTTGGGGGGTCCGGGCCCCCCCCAGGAAACACGGTGAGGCCCCAGCGCCTGCAGCCAAAGCTGG
 CACGATCTATGGGGCAGGTGCCGCTCTGCCTAGAAAAGCCAGGGGCTCTGCTGCCGTGCCCTC
 CAGAGCCACAGCGGGCAGGACTCCTCCAGCACCACCACACCCAGTGGCCCCGAGACCCCTCTG
 AGAACAGTGAGGCTGGTCCCTCGTGCCGTTCCAGCCGGTGCCCGGCCAGTGGGGAGGACACAGC
 CTAGGAACCAGCTGCCTGAGACCAGGGTGCCCTCTGGGCTGTCTCCCGCGTGGCGGAGACCCC
 AAGCACGCAGCCACCCATTTCCGGAGCTGCAGGATAGAGCTTCCTCTTGATCTCTGTTTTTAAG
 CAGAAATTCATTGTGCAGAAAAGTCCTCCAGAGCTCTGTGGCCCCGCTCGGATCCGCTGGACC
 CCCATGCCTGGCTGATCCCTGCCCACGTGGGGCAGGCCACATCTAACCCCCACAAGTCACTG
 CCTCACTGCACCTGCCAAGGCTGCCCTGGCGCTGAGTCCTGGGGTCCCTCCCGAGTTCTCTGG
 GAGAAAGGCGCCGTCGTGGCCGCTCCCGCACGCCAGGCCCGGGCTCCACCGTGGGTCTCAGA
 CGCCCTGCGGCACCGGCACCGTCTGCTTTAGCATGGGACCCCCATCTGAGGGGTGGCCTGGCC
 TTCGGGGTCCCCACGCTCCTTTGCGAAGTCCACTGTGGGTGCCATCATGGTCTCCGGGACCTG
 GGCCAGCGGGAACGTGGGGGCACTGGGTGTGCTGATATAAAGTCGGCATTACTCAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

539/615

FIGURE 534

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108765
><subunit 1 of 1, 189 aa, 1 stop
><MW: 19464, pI: 9.60, NX(S/T): 4
MALRHLALLAGLLVGVASKSMентаAQLPECCVDVVGVNASCPGASLCGP GCYRRWNADGS
ASCVRCGNGTLPAYNGSECRSFAGPGAPFPMNRSSGTPGRPHPGAPRVAASLFLGTFFIS
SGLILSVAGFFYLKRSSKLPRACYRRNKAPALQPGEAAAMIPPPQSSDVGSAGKEDPPRQ
GRPPIPAPP

Important features of the protein:

Signal peptide:

Amino acids 1-18

Transmembrane domain:

Amino acids 111-129

N-glycosylation sites:

Amino acids 38-42;68-72;75-79;92-96

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 134-138

N-myristoylation sites:

Amino acids 11-17;36-42;43-49;59-65;69-75;122-128

540/615

FIGURE 535

TGGATCTGCGGGAATGTGGGCTGGAGAGGTCTTGCCGTGGTACCAGCCTCCAGCCTGCCCCCA
GGACTGCCCCCTGACCCAGGCGCGCCCGCTGCTCGGTGGCAGGAGGGCCGGCGGAGCGCC**ATGG**
CCTGCATCCTGAAGAGAAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCGTTCCTTTTCCTGC
TGGTTGTGCGTCTTGTAATGAAGTGAATTTCCCATTTGCTACTAAACTGCTTTGGACAACCTG
GTACAAAGTGGATAACCATTCCTACACATACAGGCGGCCCTTCGAACTCACTATGGATACA
TAAATGTGAAGACACAAGAGCCTTTGCAACTGGACTGTGACCTTTGTGCCATAGTGTCAAACCT
CAGGTCAGATGGTTGGCCAGAAGGTGGGAAATGAGATAGATCGATCCTCCTGCATTTGGAGAA
TGAACAATGCCCCCACCAAAGGTTATGAAGAAGATGTCGGCCGCATGACCATGATTCGAGTTG
TGTCCCATACCAGCGTTCCTCCTTTTGCTAAAAAACCTGATTATTTTTTCAAGGAAGCGAATA
CTACTATTTATGTTATTTGGGGACCTTTCCGCAATATGAGGAAAGATGGCAATGGCATCGTTT
ACAACATGTTGAAAAAGACAGTTGGTATCTATCCGAATGCCCAAATATACGTGACCACAGAGA
AGCGCATGAGTTACTGTGATGGAGTTTTTAAGAAGGAACTGGGAAGGACAGTACAGAGT**TGAC**
CATGCAGTGTTGATTGATCGAACAGCAACCACCACATACATGTCCTGCCCCACCACAAAAGGA
AGGAAGGAATAAAAGAAAGAAAGAAAGAAACAAACAAACAAACAAACAAACTAAGCAAGACA
AAACAAATACCCATGTCAGTGGTTCAAAGATTAAGATTGTGGCTTTGTGTAAAGTTCTTTCCC
TTTGTAGACTTGCTGCATAATTATTCAGGTATGATGGTTACAGTTTTTAAAAAGGAAGGGAAA
TTGTGGTATGTGGTATGTAAATATTTTTAAATGTTGTCTCTCTGTTTTGATCAGTTTTTGTTT
TATTCAATTTGTCTTTATTAAATCTTATCAAAGCA

541/615

FIGURE 536

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108783
><subunit 1 of 1, 210 aa, 1 stop
><MW: 24022, pI: 9.51, NX(S/T): 1
MACILKRKSVIAVSFIAAFLFLLVVRLVNEVNFPLLLNCFGQPGTKWIPFSYTYRRPLRT
HYGYINVKTQEPLQLDCDLCAIVSNSGQMVGQKVGNEIDRSSCIWRMNNAPTKGYEEDVG
RMTMIRVVSHTSVPLLLKNPDYFFKEANTTIYVIWGPFRNMRKDGNGIVYNMLKKTVGIIY
PNAQIYVTTEKRMSYCDGVFKKETGKDSTE
```

Important features of the protein:

Signal peptide:

Amino acids 1-27

N-glycosylation sites:

Amino acids 148-152

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 6-10;191-195;201-205

N-myristoylation sites:

Amino acids 41-47;87-93;91-97;167-173;178-184

Downloaded from www.elsevier.com

FIGURE 537

GTTTTATTGACAATACATGCATCATATCTTTTGACTTTGAAGGATATCTCATGTCAAAGGAAT
CAAGTTATGATTTATAGAGGATTGAGCTGGAATACCTTGTGGGTGCTGGCTGAGGGTGGCAAA
ACGCCTACCGAGACATGAAGGTTTTAGCCACTAGTTTTGTCCTTGGGAGCCTGGGGTTGGCCT
TCTACCTGCCTTTGGTGGTGA CTACACCTAAAACACTGGCCATCCCTGAGAAGCTGCAAGAAG
CTGTGGGGAAAGTTATCATCAATGCCACAACCTGTACTGTACCTGTGGCCTTGGCTATAAGG
AGGAGACCGTCTGTGAGGTGGGCCCTGATGGAGTGAGAAGGAAATGTCAGACTCAGCGCTTAGAA
TGTCTGACCAACTGGATCTGTGGGATGCTCCATTTACACATTCTCATTGGCAAGGAATTTGAG
CTTAGCTGTCTGAGTTGAGACATCTTGGAGTTTGGACAGGAAGCTTTCCGGTTACCTGGAGA
CTTGCTCGAGGTGTCTCTCCACTGACGATGAGGTCTTCAAACCCCTTTCAAGCCAACTCCAC
TTTGTGAAGTTTAAATATGCTCAGGAGTATGACTCTGGGACATATCGCTGTGATGTGCAGCTG
GTAAAAA ACTTGAGACTTGTCAAGAGGCTCTATTTTGGGTTGAGGGTCCTTCCTCCTAACTTG
GTGAATCTGAATTTCCATCAGTCACTTACTGAGGATCAGAAGTTAATAGATGAGGGATTGGAA
GTTAATCTGGACAGCTACTCCAAGCCTCACCACCCAAAGTGGA AAAAGAAGGTGGCGTCAGCC
TTGGGAATAGGAATTGCCATTGGAGTGGTTGGTGGCGTGTTGGTGAGGATTGTCCTCTGTGCG
CTAAGGGGGGGCCTGCAGCAGTGA CAGCTTCAAGAACTTAACAGCCTTGCTCCTGAAGAACTG
GCTGCCCAGGAAGCCAAGCTAGCTTTTTAGGGGAGTGTTCCAGCTGCTGGTAGTGGATCAGCT
TAGAGGGAACACTCCACAGCCAAAAGAATGAGTGGGAGAAATGGAGGGGACAATCTCCTGGG
AGCTATGCGCAGTAACCTA ACTTCCCTTATGTCCCATGGATCTCTTCTGATCTTCCCTGCCCA
TTGGGTACCCAGGAACTGCAAGCATTGCCTGTGTTCTTGGAAGAGTTCTAAGAAGCTTGCA
TTCATTTTCTACCCCTTTATGACTTGGATGCCTCCCCACCTCCATTTCCCCTCTTCTGAGCTGT
GTATTCATGTAGAGGGATGTATTCAGCCTTTTTAGTGAACATTTTTTTTCAATAAAAGTAATT
CACAGTAA

543/615

FIGURE 538

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108789
><subunit 1 of 1, 255 aa, 1 stop
><MW: 28440, pI: 8.92, NX(S/T): 1
MKVLATSFVLGSLGLAFYLPLVVTTTPKTLAIPEKLQEAVGKVIINATTCTVTCGLGYKEE
TVCEVGPDGVRRKCQTQRLECLTNWICGMLHFTILIGKEFELSCLSSDILEFGQEAFT
WRLARGVISTDDEVFKPFQANSHFVKFKYAEYDSGTYRCDVQLVKNLRLVKRLYFGLRV
LPPNLVNLNFHQSLTEDQKLIDEGLEVNLDSYKPHHPKWKKKVASALGIGIAIGVVGGV
LVRIVLCALRGGLQQ

Important features of the protein:

Signal peptide:

Amino acids 1-30

Transmembrane domain:

Amino acids 225-244

N-glycosylation site:

Amino acids 45-49

N-myristoylation sites:

Amino acids 126-132;156-162;204-210;229-235;231-237;235-241

Downloaded from www.nrjonline.org

544/615

FIGURE 539

GCGCTCATCACTGGCTGGGGACAGAGCCGGGCACCAAGGAGCGACAGGATCCCGAAGAGAGAG
AGAGAAGGCAGCGAGGGAAGGAGGACCCCGGCAGGCAGCAGCATGAATTCAGCCCAGCGCAC
TACCTGCTGCCTCTCCTGCCTGCGCTGGTCCTCAGCACCAGACAGGACTATGAAGAGCTAGAA
AAGCAGCTGAAAGAAGTCTTTAAGGAGCGAAGCACCATTCTTCGTCAGCTGACAAAGACATCA
AGAGAACTTGATGGAATTAAAGTCAATCTTCAGTCCTTAAAAAACGATGAGCAGTCTGCCAAA
ACTGATGTTTCAAGAACTTCTGGAATTAGGACAGAAACAAAGAGAAGAAATGAAGTCTCTTCAG
GAGGCCCTGCAAAATCAGCTTAAGGAGACATCAGAGAAAGCAGAAAAACACCAGGCTACTATT
AATTTTTTTAAAGACTGAAGTTGAAAGAAAGAGCAAAATGATCCGAGACCTCCAGAATGAGGAT
TCAAGGAAGAGACCAAGAGATCTCCAGTGGAAGATAGTCTCCATGAGGACCATGTCAATATAC
TTATTGATGTATCTCTTAGTACCTAGAATAGTGGAGATTTATATTAGATACAAAATAAATATGT
GTGGAATTAATTAATAA

2025-09-24 10:55:01

[illegible][illegible]

547/615

FIGURE 542

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108936
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19472, pI: 5.71, NX(S/T): 0
MSPSGRLCLLTIVGLILPTRGQTLKDTTSSSSADSTIMDIQVPTRAPDAVYTELQPTSPT
PTWPADETPQPQTQTQQLEGTGDLVTDPEETHKSTKAAHPTDDTTTLSERPSPSTDVQTD
PQTLKPSGFHEDDPFFYDEHTLRKRGLLVAAVLFITGIIILTSKGKCRQLSRLCRNRCR
```

Important features of the protein:

Signal peptide:

Amino acids 1-21

Transmembrane domain:

Amino acids 147-162

Tyrosine kinase phosphorylation site:

Amino acids 45-52

N-myristoylation site:

Amino acids 146-152

2057094828001

FIGURE 543

CGGCTCGAGGTGAGAAGGAACTGCAAGAGTGGGGCAGAGAACCAGAGTGTCTCAGAGCAAAACC
TCCTCTATCTGCACATCCTGGGGACGAACCGGGCAGCCGGAGAGCTGCGGCCGGCCCAGTCCC
GCTCCGCCTTTGAAGGGTAAAACCCAAGGCGGGGCCTTGGTTCTGGCAGAAGGGACGCTATGA
CCGCAGAATTCCTCTCCCTGCTTTGCCTCGGGCTGTGTCTGGGCTACGAAGATGAGAAAAAGA
ATGAGAAACCGCCCAAGCCCTCCCTCCACGCCTGGCCCAGCTCGGTGGTTGAAGCCGAGAGCA
ATGTGACCCTGAAGTGTCTCAGGCTCATTCCTCAGAATGTGACATTTGTGCTGCGCAAGGTGAACG
ACTCTGGGTACAAGCAGGAACAGAGCTCGGCAGAAAACGAAGCTGAATTCCTTCACGGACC
TGAAGCCTAAGGATGCTGGGAGGTACTTTTGTGCCTACAAGACAACAGCCTCCCATGAGTGGT
CAGAAAGCAGTGAACACTTGCAGCTGGTGGTCTACAGATAAACACGATGAACTTGAAGCTCCCT
CAATGAAAACAGACACCAGAACCATCTTTGTGCGCATCTTCAGCTGCATCTCCATCCTTCTCC
TCTTCCTCTCAGTCTTCATCATCTACAGATGCAGCCAGCACGGTTCATCATCTGAGGAATCCA
CCAAGAGAACCAGCCATTCCAACTTCCGGAGCARGAGGCTGCCGAGGCAGATTTATCCAATA
TGGAAAGGGTATCTCTCTCGACGGCAGACCCCCAAGGAGTGACCTATGCTGAGCTAAGCACCA
GCGCCCTGTCTGAGGCAGCTTCAGACACCACCCAGGAGCCCCCAGGATCTCATGAATATGCGG
CACTGAAAGTGTAGCAAGAAGACAGCCCTGGCCACTAAAGGAGGGGGGATCGTGCTGGCCAAG
GTTATCGGAAATCTGGAGATGCAGATACTGTGTTTCCTTGCTCTTCGTCCATATCAATAAAAT
TAAGTTTCTCGTCTTA

549/615

FIGURE 544

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119510
><subunit 1 of 1, 236 aa, 1 stop
><MW: 26079, pI: 5.05, NX(S/T): 3
MTAEFLSLLCLGLCLGYEDEKKNEKPPKPSLHAWPSSVVEAESNVTLKCQAHSQNVTFVL
RKVNDSGYKQEQSSAENAEFPFTDLKPKDAGRYFCAYKTTASHEWSESSEHLQLVVTDK
HDELEAPSMKTDTRTIFVAIFSCISILLFLSVFIIYRCSQHGSSSEESTKRTSHSKLPE
QEAAEADLSNMERVSLSTADPQGVTYAELSTSALSEAASDTTQEPPGSHEYAALKV

Important features of the protein:

Signal peptide:

Amino acids 1-16

Transmembrane domain:

Amino acids 135-153

N-glycosylation sites:

Amino acids 44-48;55-59;64-68

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 171-175

Tyrosine kinase phosphorylation sites:

Amino acids 61-69;87-95

N-myristoylation sites:

Amino acids 12-18;203-209

20250404 10:43:00

FIGURE 545

GGCGGCCCCGAGCTGGGAGCGCGGGGAAGGCGGTTGGGGTTCTGACAGCTGCGCGCGATCCTG
CTCTCTCTCAGCCGCCTGTGGACATGCGCAAAGGGCCCTCTCCTGAGTCCAGATGATGCTCAT
ACCAATGGCTTTCAGTGATGGCGGTGACTGAACCGAAATGGGTCTCGGTCTGGAGCCGCTTCCT
CTGGGTGACGCTGCTGAGCATGGTGCTGGGGTCCCTGCTGGCCCTGCTGCTGCCGCTGGGGGC
TGTGGAGGAGCAGTGCTTGGCTGTGCTCAAAGGCCTCTACCTGCTCAGGAGCAAACCGGACAG
GGCGCAGCATGCCGCCACCAAGTGCAACGAGCCCGTCCACGGAGCTCAGCATCACCTCCAGGGG
CGCGACGCTGCTGGTGGCCAAGACCAAGGCCTCTCCAGCGGGTAAGTTGGAAGCCAGAGCTGC
CCTGAACCAGGCCCTGGAGATGAAGCGCCAGGGCAAGCGGGAAAAAGCCCAAAGCTCTTCAT
GCACGCCCTCAAGATGGACCCGGACTTCGTGGACGCGCTCACCGAGTTTGGCATCTTCTCGGA
AGAAGACAAGGACATCATCCAGGCGGACTACTTGTACACCAGAGCATTTGACCATCTCACCTA
CCATGAGAAAGCACTGGTCAACCGCGATCGGACACTGCCTCTTGTGGAAGAGATCGACCAGAG
GTATTTTCAGCATCATCGACAGCAAAGTGAAGAAGGTCATGTCCATCCCCAAGGGGAACTCAGC
TCTGCGCAGGGTCATGGAGGAGACCTACTACCATCACATCTACCACACAGTGGCCATCGAGGG
CAACACCCTCACCTCTCGGAAATCAGGCACATCCTGGAGACCCGCTACGCCGTGCCCGGGAA
GAGCCTGGAGGAGCAGAACGAGGTCATAGGCATGCATGCAGCCATGAAGTACATCAACACGAC
TCTGGTTTTCGCGCATCGGCTCCGTCACCATCAGCGACGTGCTGGAGATCCACAGGCGGGTGCT
GGGCTACGTGGACCCCGTGGAAGCCGGCAGGTTTCGGACAACACAGGTCCTGGTTCGGACACCA
CATCCCTCCCCATCCCGAGGATGTGGAAGAGCAGATGCAGGAGTTTGTACAGTGGCTCAACTCC
GAGGAAGCCATGAACCTGCACCCAGTGGAGTTTGCAGCCTTAGCCCATTATAAACTCGTTTAC
ATCCACCCTTTTCATTGATGGCAACGGGAGGACCTCCCGTCTGCTCATGAACCTCATCCTCATG
CAGGCGGGCTACCCGCCCATCACCATCCGCAAGGAGCAGCGGTCCGACTACTACCACGTGTTG
GAAGCTGCCAACGAGGGCGACGTGAGGCCTTTCATTTCGCTTCATCGCCAAGTGTACTGAGACC
ACCCTGGACACCCTGCTTTTGGCCACAACCTGAGTACTCGGTGGCACTGCCAGAAGCCCAACCC
AACCCTCTGGGTTCAAGGAGACGCTTCCTGTGAAGCCCTAAACCCTAGAAATCCTCAGTGACA
AAGGCTGTCCTGAGGTAGGAAA

551/615

FIGURE 546

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119517
><subunit 1 of 1, 458 aa, 1 stop
><MW: 51778, pI: 7.81, NX(S/T): 2
MMLIPMASVMVTEPKWVSVWSRFLWVTLLSMVLGSL LALLLP LGAVEEQCLAVLKGLYL
LRSKPDRAQHAATKCTSPSTELSITSRGATLLVAKTKASPAGKLEARAALNQALEMKRQG
KREKAQKLFMHALKMDPDFVDALTEFGIFSEEDKDIIQADYLYTRALTISPYHEKALVNR
DRTLPLVEEIDQRYFSIIDSKVKKVMSIPKGNSALRRVMEETYHHIYHTVAIEGNTLTL
SEIRHILETRYAVPGKSLEEQNEVIGMHAAMKYINTTLVSRIGSVTISDVLEIHRRVLGY
VDPVEAGRFRFTTQVLVGHHIPHPQDVEKQM QEFVQWLNSEEAMNLHPVEFAALAHYKLV
YIHPFIDGNGRTSRLLMNLILMQAGYPPITIRKEQRSDYYHVLEAANECDVRPFIIRFIK
CTETTLDTLLFATTEYSVALPEAQP NHSGFKETLPVKP

Important features of the protein:

Signal peptide:

Amino acids 1-46

N-glycosylation sites:

Amino acids 275-279;446-450

Tyrosine kinase phosphorylation sites:

Amino acids 216-225;217-225;244-232

N-myristoylation sites:

Amino acids 35-41;235-241;266-272;368-374

Amidation site:

Amino acids 119-123

CC-BY-NC-ND 4.0 International license

FIGURE 547

CCTCTGTCTGTGCTCCCATCCCAGGGAGTATAGGTGGAGCCTCCAGAGCCCATGGACAGGGCA
 TGCTGGGGCTGGGCCAGCCCCAGCGGTGTCTCTAAGGCACCCCTGGGATCCCCACTGAGCTGG
 CCTACTTCAGACAGCCAGGGCCCCACCCCTCTGGCCCCCTTAGTGTCCAGCTCGTGGCCCCCTTG
 GCATTTCCACAAGACGCCAAGATGGAGATTCCCATGGGGACCCAGGGCTGCTTCTCAAAGAGC
 CTCCTGCTCTCAGCCTCAATCCTGGTCCTCTGGATGCTCCAAGGCTCCCAGGCAGCTCTCTAC
 ATCCAGAAGATTCCAGAGCAGCCTCAAAGAACCAGGACCTTCTCCTGTGAGTCCAGGGTGTG
 CCAGACACCTTCCAGGACTTCAACTGGTACCTGGGGGAGGAGACGTACGGAGGCACGAGGCTA
 TTTACCTACATCCCTGGGATACAACGGCCTCAGAGGGATGGCAGTGCCATGGGACAGCGAGAC
 ATCGTGGGCTTCCCCAATGGTTCCATGCTGCTGCGCCGCGCCAGCCTACAGACAGTGGCACC
 TACCAAGTAGCCATTACCATCAACTCTGAATGGACTATGAAGGCCAAGACTGAGGTCCAGGTA
 GCTGAAAAGAATAAGGAGCTGCCAGTACACACCTGCCACCAACGCTGGGATCCTGGCGGCC
 ACCATCATTGGATCTCTTGCTGCCGGGGCCCTTCTCATCAGCTGCATTGCCTATCTCCTGGTG
 ACAAGGAAGTGGAGGGGGCCAGAGCCACAGACTGCCTGCTCCGAGGGGCCAGGGATCTCTGTCC
 ATCTTGCTGCTCGGCTGTATCCCCAGTGCCTTCAGTGACGCCCAGCACATGGATGGCGACCACA
 GAGAAGCCAGAATTGGGCCCTGCTCATGATGCTGGTGACAACAACATCTATGAAGTGATGCCC
 TCTCCAGTCCTCCTGGTGTCCCCCATCAGTGACACAAGGTCCATAAACCCAGCCCGGCCCTG
 CCCACACCCCCACACCTGCAGGCGGAGCCAGAGAACCACCAGTACCAGCAGGACCTGCTAAAC
 CCGACCCCTGCCCCCTACTGCCAGCTGGTGCCAACCTTCCTGATGGGTCTCTGGGCCAGGCCAGC
 CAGGGAGAAGACAAGGCCCCAGCCCTCCTCTGGGAGCCTCACACCTGAGACCAGCAGGACAAG
 GCCATTGGGGGCTGTGGGGCCGATGAGGTGGACTCAGCCAAAGACTCAGCAGCACATGGGGCA
 GGTGTCCTGGCAGGGGGACAGGAGACTGTAAACAGGCCCAGGTCTTGTGTCAGCCCCCTGAATGC
 ACGCCCGCCTTCGGTCTGTTCCCTTCAAGCAAGCTGGCCTGGGCCATGTGCCTGTGAAAGGCAG
 GCTCTGGCCCCCTTTCATGCCAAAGTCCCCAAGATCTGGATATCTGGGGACAAGATGGTGGC
 CTCAGGCCTGCCTCCCAGGCAGTTGGCTGGGCTCCCAACTGTCTGTCCTCAATGCCCTACCCC
 AACTCCACTAGTGACCCTCAGAGTCTTCTCCCCTTAGGACAAGGCAGACACCCCACCATGCGG
 GCCTCAGGTGGCAGAGAGGCCAGCCTCACAGGCCTGTGGCCCCACACACCAGTCCCAGCAAG
 GTGACCACGGCTGCTGGACCCCTTCCCTGTTTCAGGCAGGCCCAGCCCCTCTCAGAACCTGCTG
 CCAGCTGCTGGTCTTGGCCCCCACCCTGAATCTTACTGAGTCCCTCTGGGCAGCAGCTCCCTT
 CTCCACCCCCACCCCAGCACCCGTCCCAAATGTGGCCTCAGCTTGTCTCCCTTCCCCAACT
 ATGCATTCATTAGCAATAAATGAGCCTTTGCTGCA

553/615

FIGURE 548

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119535
><subunit 1 of 1, 300 aa, 1 stop
><MW: 32638, pI: 6.02, NX(S/T): 1
MEIPMGTQGCFSKSLLLSASILVLWMLQGSQAALYIQKIPEQPQKNQDLLLSVQGVPDF
QDFNWYLGEETYGGTRLFTYIPGIQRPQRDGSAMGQRDIVGFNGSM LLRRAQPTDSGT
YQVAITINSEWTMKAKTEVQVAEKNKELPSTHLP TNAGILAATIIGSLAAGALLISCIAY
LVTRNWRGQSHRLPAPRGQGSLSILCSAVSPVPSVTPSTW MATTEKPELGP AHDAGDNNI
YEVMPSPVLLVSPISDTRSINPARPLPTPPHLQAEPENHQYQQDLLNPD PAPYCQLVPTS
```

Important features of the protein:

Signal peptide:

Amino acids 1-32

Transmembrane domain:

Amino acids 159-178

N-glycosylation site:

Amino acids 104-108

N-myristoylation sites:

Amino acids 6-12;29-35;55-61;91-97;157-163;165-171

protein sequence

FIGURE 549

GCCACTCACACCATCTGCTAATGGGACAGCTCACTCTTCCCTCCAAACCATGGCCTTGGCTCA
 AGAGCTTCCTTGTTTCTGGAATGTTCTTTCTCCAGCTCCAGGTGTTGAAATTTCTGCCTGGTC
 TGGGTCTCCTGTTGAAGGACGCCCTCCACTGGGAAGGATCCTCTTGCCTTCACCACTTGTCTT
 CACCAGCCCCCTGCTCCCCCTCTTCTCTTGGGGCTGTTGTCTGTTGATACTTTTTTTTTTGT
 GTGTTTGACACACATCTTCTTCTCACCTCTAACACAGTTCTCAACCACAGCACTTTTGTCCC
 TGGAGATGTTGGCAGTGTCCAGAGGCGTGTTGATGGTCCCCTGGGGTGGGGGTGCTGCTGGCA
 CCAGATGGTAGGGAGATGCCAGGGGTGCTGCTCCACACCCTATGGGACACTGCACAGTACACC
 TGGCCTGTGTCCCCACAGCGAGAGCTGGCCCTGGGCAGGCGTGGTCCCTGCGGTGTGTGTTG
 GTTGGGATCCTCCACAGTGACAGACGGTGCGCTCTGCCCACGTTTCCACACAGCTCTTTTGCT
 TGTGGAGCTCACCCCTTTGCAGAGAGCTCATTTCCCTGCGGTCTTTGGCCTGCAGAAATAAAA
TGAGGGGGTGGTGAATTACACCCCTGCTGGTTACACATGGAAAACCTCAGGAGTGAGAAATTTTGT
 GGAGAGCAAGAGAGGTGAGACTGGGGTGTGGCTGCCAGCCAGGCGGTCCCTCAGCCCCCTGGA
 GAAGCGGGGTGGGGCCTGCACACCGAGTCTTCCAGTGAGTCCAGTGATGCTCTCTCCTCTTC
 CTCCCAGTCACCTTTCTCTCCAGTGCCACTACTGCGCTTTCGATGCAGAATAATTCAGTATTT
 GCGGACTTGAAGTCGGACGAGATGGAGCTGCTCTACTCAGCCTACGGAGATGAGACAGGCGTG
 CAGTGTGCGCTGAGCCTGCAGGAGTTTGTGAAGGATGCTGGGAGCTACAGCAAGAAAGTGGTG
 GACGACCTCCTGGACCAGATCACAGGCGGAGACCACTCTAGGACGCTCTTCCAGCTGAAGCAG
 AGAAGAAATGTTCCCATGAAGCCTCCAGATGAAGCCAAGGTTGGGGACACCCCTAGGAGACAGC
 AGCAGCTCTGTTCTGGAGTTCATGTCTGATGAAGTCTATCCCGACGTTTCTGTGGATATCTCC
 ATGCTCAGCTCTCTGGGGAAGGTGAAGAAGGAGCTGGACCCCTGACGACAGCCATTTGAACTTG
 GATGAGACGACGAAGCTCCTGCAGGACCTGCACGAAGCACAGGCGGAGCGCGGCGGCTCTCGG
 CCGTCGTCCAACCTCAGCTCCCTGTCCAACGCCTCCGAGAGGGACCAGCACCACCTGGGAAGC
 CCTTCTCGCCTGAGTGTGCGGGAGCAGCCAGACGTCACCCACGACCCCTATGAGTTTCTTCAG
 TCTCCAGAGCCTGCGGCCTCTGCCAAGACCTAACTCTAGACCACCTTCAGCTCTTTTATTTTA
 TTTTTTTAGTTTTTATTTTGCACGTGTAGAGTTTTTGTGTCATCAGACAAGGACTTTTGATCCTGTC
 CCCTTTGGCATGCGGGAAGCAGCCGCGGGGAGGTAATGAATTGTCTGTGGTATCATGTCAGCA
 GAGTCTCCAAGCCCCACGAACCCTGAGGAGTGGAGTCATACGCGAAGGCCATATGGCCATCGT
 GTCAGCAGAGAGAGTCTCTGTACACAGCCCCGTGAACCCTGAGGAGTGGAGTCATACACGAAG
 GCGTGTGGCCATCGTGTACAGCAGAGAGTCTCTGTACACAGCCCCGTGAACCCTGAGGAGTGG
 AGTCATACGCGAAGGGTGTGTGGCCAGGCTGCAGAGCTGCGTGCCGTTTTGTGTCCGAGCATCA
 CGTGTGGCTCCAGCCCTTGTTTCTGCCAGTGTAGACACCTCTGTCTGCCCCACTGTCTTGGGG
 TCGCTCTTGGGAGGCACAGGCATGGGTGTGTCTGGCCTCATTTCTGTATCAGTCCAGTGTGTTT
 CTGTCATAGTTTTGTGTCTCCAGGCAGGCCATGGTAGGGGCCTCGCAGGGGCCATTGGGGAGC
 ACAGGGCCAGGCTGGGGTGAGGAGAGCTCCCCTGTTTTCTGTTTAATTGATGAGCCTGGGAAA
 GGAGTGTGTTCTGCCTGCCCGTTACAGTGAGCGTTCGTTGCCATAAAACGTTTTTCTAACTG
 GGAA

555/615

FIGURE 550

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119537
><subunit 1 of 1, 104 aa, 1 stop
><MW: 11136, pI: 8.20, NX(S/T): 0
MLAVSRGVLMVPLGLGVLLAPDGREMPGVLLHTLWDTAQYTWPVSPPTARAGPGQAWSLRC
VLVGILHSDRRRCALPTFPHSSFACGAHPFAESSFPCGLWPAEVK
```

Important features of the protein:

Signal peptide:

Amino acids 1-20

N-myristoylation sites:

Amino acids 53-59;64-70;97-103

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 74-85

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 551

CGCCCTTAGCATGCGGTTTCTGCCGAGCGCCCTGCACCCGCTGTCTCTCCTGGTGAGGCCATCATGCTGGCCAT
 GACCCCTGGCCCTGGGTACCTTGCCTGCCTTCTTACCCTGTGAGCTCCAGCCCCACGGCCTGGTGAACCTGCAACTG
 GCTGTTCCCTGAAGTCTGTGCCCCACTTCTCCATGGCAGCACCCCGTGGCAATGTCACCAGCCTTTCCCTTGTCTC
 CAACCGCATCCACCACCTCCATGATTCTGACTTTGCCACCTGCCAGCCTGCGGCATCTCAACCTCAAGTGGA
 CTGCCCGCCGTTGGCCTCAGCCCCATGCACTTCCCTGCCACATGACCATCGAGCCAGCACCTTCTTGGCTGT
 GCCCCACCTGGAAGAGCTAAACCTGAGCTACAACAACATCATGACTGTGCTGCGCTGCCCAAATCCCTCATATC
 CCTGTCCCTCAGCCATACCAACATCCTGATGCTAGACTCTGCCAGCCTCGCCGGCCTGCATGCCCTGCGCTTCT
 ATTATGAGGACGGCAACTGTTATTACAAGAACCCTGCAGGCAGGCACTGGAGGTGGCCCCGGGTGCCCTCCTTGG
 CCTGGGCAGCCTCACCCACCTGTCACTCAAGTACAACAACCTCACTGTGGTGCCCCGCAACCTGCCTTCCAGCCT
 GGAGTATCTGCTGTTGTCTTACAACCGCATCGTCAAACCTGGCGCCTGAGGACCTGGCCAATCTGACCGCCCTGCG
 TGTGCTCGATGTGGGCGGAAATTGCCGCCGTGCGACCACGCTCCCAACCCCTGCATGGAGTGCCCTCGTCACTT
 CCCCCAGCTACATCCCGATACCTTCAGCCACCTGAGCCGTCTTGAAGGCCTGGTGTGAAGGACAGTTCTCTCTC
 CTGGCTGAATGCCAGTTGGTTCCGTGGGCTGGGAAACCTCCGAGTGCTGGACCTGAGTGAGAACTTCTCTACAA
 ATGCATCACTAAAACCAAGGCCCTCCAGGGCCTAACACAGCTGCGCAAGCTTAACCTGTCTTCAATTACCAAAA
 GAGGGTGTCTTTTGCCACCTGTCTGTGGCCCTTCTTTCGGGAGCCTGGTCTGCCCTGAAGGAGCTGGACATGCA
 CGGCATCTTCTTCCGCTCACTCGATGAGACCACGCTCCCGGCCACTGGCCCCGCTGCCCATGCTCCAGACTCTGCG
 TCTGCAGATGAACCTCATCAACCAGGCCAGCTCGGCATCTTCAGGGCCTTCCCTGGCCTGCGCTACGTGGACCT
 GTCGGACAACCGCATCAGCGGAGCTTCGGAGCTGACAGCCACCATGGGGGAGGCAGATGGAGGGGAGAAGGTCTG
 GCTGCAGCCTGGGGACCTTGCTCCGGCCCCAGTGGACACTCCAGCTCTGAAGACTTCAGGCCCAACTGCAGCAC
 CCTCAACTTCACCTTGGATCTGTACGGAACAACCTGGTGACCGTGACCCGAGATGTTTGGCCAGCTCTCGCA
 CCTGCAGTGCCTGCGCCTGAGCCACAACCTGCATCTCGCAGGCAGTCAATGGCTCCAGTTCTTCCGCTGACCCG
 TCTGCAGGTGCTAGACCTGTCCACAATAAGCTGGACCTTACCACGAGCACTCATTACCGGAGCTACCACGACT
 GGAGGCCCTGGACCTCAGCTACAACAGCCAGCCCTTTGGCATGCAGGGCGTGGGCCACAACCTCAGCTTCGTGGC
 TCACCTGCGCACCCCTGCGCCACCTCAGCCTGGCCCCACAACAACATCCACAGCCAAGTGTCCCAGCAGCTCTGCAG
 TACGTGCTGCGGGCCCTGGACTTCAGCGGCAATGCACTGGGCCATATGTGGGCCGAGGGAGACCTCTATCTGCA
 CTTCTTCCAAGGCCTGAGCGGTTTGATCTGGCTGGACTTGTCCAGAACCCTGCACACCCTCCTGCCCAAAC
 CCTGCGCAACCTCCCCAAGAGCCTACAGGTGCTGCGTCTCCGTGACAATTACCTGGCCTTCTTTAAGTGGTGGAG
 CCTCCACTTCTTCCCAAACCTGGAAGTCTCGACCTGGCAGGAAACCAGCTGAAGGCCCTGACCAATGGCAGCCT
 GCCTGCTGGCACCCGGCTCCGGAGGCTGGATGTGAGCTGCAACAGCATCAGCTTCGTGGCCCCCGGCTTCTTTTC
 CAAGGCCAAGGAGCTGCGAGAGCTCAACCTTAGCGCCAACGCCCTCAAGACAGTGGACCACCTCTGGTTTGGGCC
 CCTGGCGAGTGCCCTGCAATACTAGATGTAAGCGCCAACCCCTCTGCACTGCGCCTGTGGGGCGGCCCTTTATGGA
 CTTCTGCTGGAGGTGCAGGCTGCCGTGCCCGTCTGCCAGCCGGGTGAAGTGTGGCAGTCCGGGCCAGCTCCA
 GGGCCTCAGCATCTTTGACAGGACCTGCGCCTCTGCCCTGGATGAGGCCCTCTCCTGGGACTGTTTCGCCCTCTC
 GCTGCTGGCTGTGGCTCTGGGCCTGGGTGTGCCCATGCTGCATCACCTCTGTGGCTGGGACCTCTGGTACTGCTT
 CCACCTGTGCCCTGGCCTGGCTTCCCTGGCGGGGGCGGCAAAGTGGGCGAGATGAGGATGCCCTGCCCTACGATGC
 CTTCTGTGGTCTTCGACAAAACGCAGAGCGCAGTGGCAGACTGGGTGTACAACGAGCTTCGGGGGCAGCTGGAGGA
 GTGCCGTGGGCGCTGGGCACTCCGCCCTGTGCCTGGAGGAACGCGACTGGCTGCCTGGCAAAACCCCTCTTTGAGAA
 CCTGTGGGCCTCGGTCTATGGCAGCCGCAAGACGCTGTTTGTGCTGGCCACACGGACCGGTCAGTGGTCTCTT
 GCGCGCCAGCTTCTGCTGGCCAGCAGCGCCTGCTGGAGGACCGCAAGGACGTCTGTGGTGTGGTGATCCTGAG
 CCCTGACGGCCGCGCTCCCGCTACGTGCGGCTGCGCCAGCGCCTCTGCCGCCAGAGTGTCTCTCTGGCCCCA
 CCAGCCCAGTGGTCAGCGCAGCTTCTGGGCCAGCTGGGCATGGCCCTGACCAGGGACAACCACCACTTCTATAA
 CCGGAACCTTCTGCCAGGGACCCACGCCGAATAGCCGTGAGCCGGAATCCTGCACGGTGCCACCTC

557/615

FIGURE 552

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119714
><subunit 1 of 1, 1032 aa, 1 stop
><MW: 115799, pI: 8.61, NX(S/T): 12
MGFCRSALHPLSLLVQAIMLAMTLALGTLPAFLPCELQPHGLVNCNWLFLKSVPHFMSMAA
PRGNVTSLSLSSNRIHHLHDSDFAHLPRLRLNLKWNCPVGLSPMHFPCCHMTIEPSTFL
AVPTLEELNLSYNNIMTVPALPKSLISLSLSHTNIMLDSASLAGLHALRFLFMDGNCY
KNPCRQALEVAPGALLGLGSLTHLSLKYNNLTVPVPRNLPSSLEYLLLSYNRIVKLAPEDL
ANLTALRVLDVGGNCRRCDHAPNPCMECPRHFPQLHPDTFSHLSRLEGLVLKDSSLSWL
ASWFRGLGNLRVLDLSENFYKCIITKTKALQGLTQLRKLNLNFNYQKRVSFAHLSLAPSF
GSLVALKELDMHGIFFRSLDETTLRLPLARLPLQLTLRLQMNFINQAQLGIFRAFPGLRYV
DLSDNRISGASELTATMGEADGGEKVWLQPGDLAPAPVDTPSSEDFRPNCSTLNFTLDLS
RNNLVTVQPEMFAQLSHLQCLRLSHNCISQAVNGSQFLPLTGLQVLDLSHNKLDLYHEHS
FTELPRLEALDLSYNSQPFQMVGHNFSFVAHLRTLRLHLSLAHNNIHSQVSQQLCSTSL
RALDFSGNALGHMWAEGDLYLHFFQGLSGLIWLDSLQNRLHTLLPQTLRLNLPKSLQVLR
RDNYLAFFKWWSLHFLPKLEVLDLAGNQLKALTNGSLPAGTRLRRLDVSCNSISFVAPGF
FSKAKELREINLSANALKTVDHWSFGPLASALQILDVSANPLHCACGAAFMDFLLEVQAA
VPGLPSRVKCGSPGQLQGLSIFAQDLRLCLDEALSWDCFALSLLAVALGLGVPMHLHLCG
WDLWYCFHLCLAWLPWRGRQSGRDEDALPYDAFVVFDKTQSAVADWVYNELRGQLEECRG
RWALRLCLEERDWLPKTLFENLWASVYGSRKTLFVLAHTDRVSGLLRASFLLAQQRLL
DRKDVVVLVILSPDGRRSRYVRLRQRLCRQSVLLWPHQPSGQRSFWAQLGMALTRDNH
YNRNFCQGPTAE

Important features of the protein:

Signal peptide:

Amino acids 1-30

Transmembrane domain:

Amino acids 818-835

N-glycosylation sites:

Amino acids 64-68;129-133;210-214;242-246;300-304;340-344;
469-473;474-477;513-517;567-571;694-698;731-735

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 347-351

Tyrosine kinase phosphorylation site:

Amino acids 863-871

N-myristoylation sites:

Amino acids 27-33;41-47;63-69;193-199;361-367;409-415;
563-569;607-613;695-701;794-800;929-935;945-951;
1010-1016

Amidation site:

Amino Acids 974-978

Leucine zipper patterns:

Amino acids 204-226;644-666;814-836

FIGURE 553

GGCGTGGGACGTGCTGCGGCGTCCTAGCTGGCTTACAGGGCGGCGGCGGGGTGTGTGTCCTCT
GTTAAGAGTGCTACTCGCCCGGGGTTGATCTGTGCATGCCACTCCTGGGTGAGACGGTGAGGT
CGGCGTCTGCGAGGACGCGGCGGTGGAGTAGAAGGGCAGCCGGAGACAGGCCCCGGCGCCCCCTT
CCGAGGCTAGACGGCCCCAGCTTCGCGGGGATC**ATGGC**ATTGCTGGTGGACCGAGTGCGGGGC
CACTGGCGAATCGCCGCGGGGCTCCTGTTCAACCTGCTGGTGTCCATCTGCATTGTGTTCCCTC
AACAAATGGATTTATGTGTACCACGGCTTCCCCAACATGAGCCTGACCCTGGTGCACCTTCGTG
GTCACCTGGCTGGGCTTGTATATCTGCCAGAAGCTGGACATCTTTGCCCCCAAAGTCTGCCG
CCCTCCAGGCTCCTCCTCCTGGCCCTCAGCTTCTGTGGCTTTGTGGTCTTCACTAACCTTTCT
CTGCAGAACCAACACCATAGGCACCTATCAGCTGGCCAAGGCCATGACCACGCCGGTGATCATA
GCCATCCAGACCTTCTGCTACCAGAAAACCTTCTCCACCAGAATCCAGCTCACGCTGATTCCCT
ATAACTTTAGGTGTAATCCTAAATTCTTATTACGATGTGAAGTTTAATTTCTTGGAATGGTG
TTTGCTGCTCTTGGTGTTTTAGTTACATCCCTTTATCAAGTGTGGGTAGGAGCCAAACAGCAT
GAATTACAAGTGAACCTCAATGCAGCTGCTGTACTACCAGGCTCCGATGTCATCTGCCATGTTG
CTGGTTGCTGTGCCCTTCTTTGAGCCAGTGTTTGGAGAAGGAGGAATATTTGGTCCCTGGTCA
GTTTCTGCTTTGCTTATGGTGCTGCTATCTGGAGTAATAGCTTTTCATGGTGAACCTTATCAATT
TATTGGATCATTGGGAACACTTCACCTGTCACCTATAACATGTTCCGACACTTCAAGTTCTGC
ATTACTTTATTTCGGAGGATATGTTTTATTTAAGGATCCACTGTCCATTAATCAGGCCCTTGGC
ATTTTATGTACATTATTTGGCATTCTCGCCTATACCCACTTTAAGCTCAGTGAACAGGAAGGA
AGTAGGAGTAAACTGGCACAACGTCCT**TAA**TGGGTTTTTTGTGGAGAAAAGAATGTTGTCCCA
AGAAGATAAAAAATATTGTTAAGTGTGCAAGTTATTA

559/615

FIGURE 554

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125170
><subunit 1 of 1, 313 aa, 1 stop
><MW: 35066, pI: 9.39, NX(S/T): 5
MALLVDRVRGHWRIAAGLLFNLLVSICIVFLNKWIYVYHGFPNMSLTLVHFVVTWLGLYI
CQKLDIFAPKSLPPSRLLLLLSFCGFVVFTNLSLQNNITIGTYQLAKAMTTPVIIAIQTF
CYQKTFSTRIQLTLIPITLGVILNSYYDVKNFLGMVFAALGVLVTSLYQVWVGAKQHEL
QVNSMQLLYYQAPMSSAMLLVAVPFFEPVFGEGGIFGPWSVSALLMVLLSGVIAFMVNLS
IYWIIGNTSPVTYNMFGHFKFCITLFGGYVLFKDPLSINQALGILCTLFGILAYTHFKLS
EQEGSRSKLAQRP

Important features of the protein:

Signal peptide:

Amino acids 1-27

Transmembrane domains:

Amino acids 46-60;75-90;153-167;192-208;221-237

N-glycosylation sites:

Amino acids 43-47;92-96;97-101;238-242

N-myristoylation sites:

Amino acids 17-23;57-63;140-146;155-161;162-168;283-289

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 555

GTTAGGCAGAGCCAAGGTGGTTGCAGACCTGGAATCAGAACAGCTTTTAGACCAACCTGAAAG
CAGGAATGTAAGCACTGTTACAGAGATTTTCGTCTTTGGCTTATTGTGCCTGCAGAGTCTAG
TGCTTCTTTGCCAGCTGTGCTGACTCAGCACTCCATGCCTGTCTTTCTGGAACCAGTCCCTGGA
GCTGGGCCATGTTTTGATTGACAGTGTGGAGCTAGCCCAGCAAGTACTCTACATGCAACCCCC
CACCCAGGCACTACCTCTGCTCCTCCTCCATGGCCTCCTGCTACACCGGCAGCTCTATGGAAC
AAGGCTGCAGGCACACAGGGGGCGCTGGAGTCAAGTGAAGTCTAACCAGGTTCTTCAGACCCA
AGACCAGCTGTGGGCAAGTCTTAGCAATCCCCGTGCTGCCATGCAAGAGCTGGCTGCTTCAGT
TTTCTACGGGGGTCTCTGCGGGGACACTGAGGACAGGGAGGCCCTGATTAGCCTCACACAAGC
CTGCCTGAGCCCCAGTAGTGGGAGCTGGGTCCAGCCACACACACCTCAGTCTTTGCTGGCCAC
GCTCATGCCCCCTCCAGCTAAGGGAGCTGGATGCAATGGCAGAGTGCAAGGCCAGATGCACC
TACTGCCCTCACCACTTGAACCCCCGGCTCTGCGGACTGAGTGAGGGCCCCCAAGCCTGGCTGT
TGCGACGCCAGAGTCGCGCTCTCTTGAGTGCCTGTCAGCGGAGTTACCCCGTGTGGGTCTCTG
AGTCTCGAAGAGGCGCCAGCTTGCGGAAAGGCGACTGCGGCAACGCCTAGTGCAAGTCAACCG
GAGGCTGGAGTCACTGCAGGATCTGCTGACCCACGTGATTGCGCAAGACGAGTCCGACGCCCC
GTGGTCAGTGCTGGGGCCAAATGCACGGCGGCCTCTGGAGGGCGTCTTAGAGACCGAGGCTCT
AGAACTGAGCCAGTTGGTGGGCACGCTACAACGCGACCTTGATTGCCTGTTGCAGCAGCTGAA
GGGCGCACCCCCGTGCCCCCTCCCGCCGCTGTGCTGCGGTGGCCACGCTCTCTGGACTGGCCG
CCTACCCTTGCCCTTGCGGACCTCATGCGCCGGCCGGTCCGCAGCCGCCCTGGCACTGGCTGCG
ACAGTTGTGCGCGCCGTGGGCAACTGTTGGTTCGTTACTTGGGCGTGGGCGCGGACGCGAGCAG
TGATGTACCAGAGCGCGTCTTCCACCTGTCAGCCTTTCGCCACCCGCGCCGCTGCTGCTGGC
ATTGCGTGGGGAAGCTGCCCTGGACCAGAATGTGCCAGCTCGAATTTCCCTGGTAGCCGAGG
CTCGGTCTCCAGTCAGCTCCAGTATAAACGTCTGGAGATGAACAGCAACCCTCTGCACTTCAG
GGTGGAGAATGGTCCAAATCCCACGGTTCAGAGAGAGGGCTGCTGCTGATCGGGCTACAGGT
CCTACATGCGGAGTGGGACCCAATAGCTGGAGCCTTGACAGGACAGTCCTTCCAGCCAACCCAG
CCCTCTGCCTCCCGTCAGCATCAGCACACAGGCCCCGGGCACCAGTGACCTGCCAGCCCCAGC
CGACCTGACTGTGTACTCGTGTCTGTGTACATGGGAGGGCCCCCTTGGCACCGCTAAGCTGCA
GAGCAGGAACATCGTGATGCATCTGCCTTTACCCACCAAGCTCACCCCCAACACCTGTGTCCA
AAGGAGGGTCCATGTGTGCAGCCCACCCCTGTCTTGAGCCCGTCTACCAAATAAAGTTGTAG
TGATTCCA

561/615

FIGURE 556

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129594
><subunit 1 of 1, 162 aa, 1 stop
><MW: 17598, pI: 6.58, NX(S/T): 1
MPVFWNQSLLELGHVLIIDSVELAQQVLYMQPPTQALPLLLHGLLLHRQLYGTRLQAHGR
WSQVTLTQVLQTQDQLWASLSNPRAAMQELAASVFYGGPLGDTEDREALISLTQACLSPS
SGSWVQPHTPQSLLATLMPLPAKGAGCNGRVQGPDPAPTALTT
```

Important features of the protein:

Signal peptide:

Amino acids 1-45

N-glycosylation site:

Amino acids 6-10

N-myristoylation sites:

Amino acids 97-103;144-150

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 557

GACCTTGAGCCCTCGAAAGCGACATGGCGGTTCTCTTAAAGCTGGGCGTTCTCTGCAGTGGCC
AAGGAGCTCGAGCTCTCCTACTCCGAAGCCGGGTGGTCAGACCCGCTTATGTGTCAGCATTTC
TCCAGGACCAGCCTACCCAAGGACGGTGTGGTACCCAGCACATTACCTGTACCAAGCCACC
ACTCTGGTTCCAAGGCTGCATCTCTCCACTGGACCAGTGAGAGGGTTGTGAGTGTCTGTCTCT
TGGGGCTGATCCCTGCTGGGTACTTGAATCCCTGCTCTGTGGTGGACTACTCTCTGGCTGCAG
CCCTCACCCCTGCACAGTCACTGGGGCCTTGGACAAGTGGTTACCGACTACGTTTCATGGGGACA
CCCTGCCGAAGGCTGCCAGGGCAGGCCTCTTGGCACTCTCAGCTTTGACCTTTGCTGGGCTTTGC
TACTTCAATTACCACGATGTGCGCATCTGCAGAGCGGTTGCCATGCTGTGGAAGCTCTGA CCT
GGGTGCAGCACTTTGATTGTGTGCCTCCTTGCCTCTGCTTTACCAATGCCGTTACCTCGCAG
TGAGGGGGGATGAAGGATAAGCCCATTGGTGGGCAGAATGTCTTCTAATTACATGGTTATTTT
CAGAATTTATTTGTTGAGGAAGAGGTTTGAGGAGTTAGGTTTCGACCATTCGTGAGTCTGTGTT
CCATACTCCACTGAGTGTGGGCACTAGCTCACAGCCTCGCGGTGAGACTGAACATTTTCATGAG
CTCATGTTGCCTTTGACCACCATTTCTTAAGGAGAGCCAGCTGATTGCTGTCAGGATAAGAGC
ATCTCTTCAGCCAGGAGGGAGGCCTGTTCCCTCCTGAGTTAGACTTTGCATGAAGCTCGAAAG
TATTCCCTTTGGAACCTCCCATTTCTTGTTTCAGGTGACACCAGCTCTGTTGATGGCTCTGCTTC
TAGGGAACATTTAATCAGGAGATGCTCTCAATGACTAATTTGTCTAAGTCTTAGGAAGGAGGT
TGAGGAAAGCTGGATTTAGACAAGTTCAATTTAGGGAGTTCTCCTTGTTTGTGGATTAAATA
TGACAGATTGCAAACAGACTACTCTTCAAATGTATCTCAATTGTGCAGAAGTGAGCTGTCCAA
AAGTATAAGACTAAGTGATAAACTGTCTTCCCACCGTGGGAGTTGTTAATGAGAAAGAAAGTG
TACTCTGAAAAACAAGGGGG

563/615

FIGURE 558

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129793
><subunit 1 of 1, 159 aa, 1 stop
><MW: 17014, pI: 9.38, NX(S/T): 0
MAVLLKLGVLCSGQGARALLRSRVVRPAYVSAFLQDQPTQGRCGTQHIHLSPSHHSGSK
AASLHWTSEVVSVLLGLIPAGYLNPCSVVDYSLAAALTLHSHWGLGQVVTDYVHGDTL
PKAARAGLLALSALTTFAGLCYFNYHDVGICRAVAMLWKL
```

Important features of the protein:

Signal peptide:

Amino acids 1-15

Transmembrane domains:

Amino acids 71-88;126-140

Glycosaminoglycan attachment site:

Amino acids 12-16

N-myristoylation sites

Amino acids 8-14;58-64;78-84;108-114;148-154

FIGURE 559

CCCAGCCCCGCGTTTCGGCTGCTCTCGAGGAGGCCGGAGTCCCCGGAGACGATGCGCCCCGCGC
 AGCCGCCTGCGCCTGCGGGAGCCGGCTGCCCTTGAGATGGAGTTGCTGCCTCTTTGGCTCTGC
 CTGGGTTTTTCACTTCCTGACCGTGGGCTGGAGGAACAGAAGCGGAACAGCCACAGCAGCCTCC
 CAAGGAGTCTGCAAGTTGGTGGGTGGAGCCGCTGACTGCCGAGGGCAGAGCCTCGCTTCGGTG
 CCCAGCAGCCTCCCCCCCCACGCCCCGATGCTCACCTGGATGCCAACCCTCTCAAGACCCTG
 TGGAATCACTCCCTCCAGCCTTACCCTCTCCTGGAGAGCCTCAGCCTGCACAGCTGCCACCTG
 GAGCGCATCAGCCGCGGCGCCTTCCAGGAGCAAGGTCACCTGCGCAGCCTGGTCCTGGGGGAC
 AACTGCCTCTCAGAGAACTACGAAGAGACGGCAGCCGCCCTCCACGCCCTGCCGGGCCTGCGG
 AGGCTGGACTTGTGAGGAAACGCCCTGACGGAGGACATGGCAGCGCTCATGCTCCAGAACCTC
 TCCTCGCTGCGGTCCGTGTCCCTGGCGGGGAACACCATCATGCGGCTGGACGACTCCGTCTTC
 GAGGGCCTGGAGCGTCTCCGGGAGCTGGATCTGCAGAGGAACTACATCTTCGAGATCGAGGGC
 GGCGCTTTCGACGGCCTGGCTGAGCTGAGGCACCTCAACCTGGCCTTCAACAACCTCCCCTGC
 ATCGTGGACTTCGGGCTCACGCGGCTGCGGGTCTCAACGTCAGCTACAACGTCCTGGAGTGG
 TTCCTCGCGACCGGGGGAGAGGCTGCCTTCGAGCTGGAGACGCTGGACCTGTCTCACAACCAG
 CTGCTGTTCTTCCCGCTGCTGCCCCAGTACAGCAAGTTGCGGACCCTCCTGCTGCGCGACAAC
 AACATGGGCTTCTACCGGGACCTGTACAACACCTCGTCGCCGAGGGAGATGGTGGCCCAGTTC
 CTCCTCGTGGACGGCAACGTGACCAACATCACCAACGTCAGCCTCTGGGAAGAATTCTCCTCC
 AGCGACCTCGCAGATCTCCGCTTCCTGGACATGAGCCAGAACCAGTTCAGTACCTGCCAGAC
 GGCTTCCTGAGGAAAATGCCTTCCCTCTCCCACCTGAACCTCCACCAGAATTGCCTGATGACG
 CTTACATTCGGGAGCACGAGCCCCCGGAGCGCTCACCGAGCTGGACCTGAGCCACAACCAG
 CTGTGCGAGCTGCACCTGGCTCCGGGGCTGGCCAGCTGCCTGGGCAGCCTGCGCTTGTTCAAC
 CTGAGCTCCAACCAGCTCCTGGGCGTCCCCCTGGCCTCTTCGCCAATGCTAGGAACATCACTAC
 ACTTGACATGAGCCACAATCAGATCTCACTTTGTCCCCTGCCAGCTGCCTCGGACCGGGTGGG
 CCCCCCTAGCTGTGTGGATTTAGGAATATGGCATCTTTAAGGAGCCTGTCTCTGGAGGGCTG
 TGGCCTGGGGGCATTGCCAGACTGCCCATTCGAAGGGACCTCCCTGACCTACTTAGACCTCTC
 AAGCAACTGGGGGGTTCTGAATGGGAGCCTCGCCCCACTCCAGGATGTTGCCCCCATGTTACA
 GGTCTGTCTCTCAGGAACATGGGCCTCCACTCCAGCTTTATGGCGTTGGACTTCTCTGGGTT
 TGGGAATCTCAGGGACTTAGATCTGTGCGGGGAATTGCTTGACCACCTTCCCAAGGTTTGGGGG
 CAGCCTGGCCCTGGAGACCCTGGATCTCCGTAGAAACTCGCTCACAGCCCTTCCCCAGAAGGC
 TGTGTCTGAGCAGCTCTCGAGAGGTCTGCGGACCATCTACCTCAGTCAGAATCCATATGACTG
 CTGTGGGGTGGATGGCTGGGGGGCCCTGCAGCATGGGCAGACGGTGGCCGACTGGGCCATGGT
 CACCTGCAACCTCTCCTCCAAGATCATCCGCGTGACGGAGCTGCCCGGAGGTGTGCCTCGGGA
 CTGCAAGTGGGAGCGGCTGGACCTGGGCCTGCTCTACCTCGTGCTCATCCTCCCCAGCTGCCT
 CACCCTGCTGGTGGCCTGCACTGTCTCGTCCTCACTTTTAAGAAGCCTCTGCTTCAGGTCAT
 CAAGAGCCGCTGCCACTGGTCCTCCGTTTACTGAACCTGGCTGTGTGCCAAGACTCGAAATTCG
 GTCCGCACACAACAGGACACTTTCTCTGCCAGCTTTCAAGATGTGATGCAGAGGCCAAGTCTG
 ACGAATTGAAGTTTCAATTAAAATTTAATATGTTTCCATTCTCATCGCCACCCACCCCCG
 CCCCCACCACCGCCAAGTCTTTTTTCCATCATTATAATTATCCTTATTATCTTGGTAAAAT
 ATTTATTAAGTGACTTTTTTCAGAAATAAAAGGCAACGTGTCTCATAAATATTTTTTAAAAAA
 AAAAAAAAAAAAAA

565/615

FIGURE 560

><subunit 1 of 1, 692 aa, 1 stop
><MW: 76366, pI: 6.07, NX(S/T): 11
MELLPLWLCLGFHFLTVGWRNRS GTATAASQGVCKLVGGAADCRGQSLASVPSSLPPHAR
MLTLTDANPLKTLWNHSLQPYPLLESLSLH SCHLERISRGA FQEQGHLRSLVLGDNCLSEN
YEETAAALHALPGLRRLDLSGNALTEDMAALMLQNLSSLSRSVSLAGNTIMRLDDSVFEGL
ERLRELDLQRNYIFEIEGGA F DGLAELRHLNLA FNNLPCIVDFGLTRLRLVLNVSYNVLEW
FLATGGEAAFELETLDLSHNQLLFFPLLPQY SKLRTL LLLRDNNMGFYRDLYNTSSPREMV
AQFLLVDGNVTNITT VSLWEEFSSSDLADLRFLDMSQNFQYLPDGFLRKMPSLSHLNLH
QNCLMTLHIREHEPPGALTELDLSHNQLSELHLAPGLASCLGSLRLFNLSNQLLGVP PG
LFANARNITTLDMSHNQISLCPLPAASDRVGPPSCVD FRNMA SLRSLSLEGCGLGALPDC
PFQGTSLTYLDLSSNWGVLNGSLAPLQDVAPMLQVLSLRNMGLHSSFMALDFSGFGNLRD
LDLSGNCLTTFFPRFGGSLALETLDLRNSLTALPQKAVSEQLSRGLRTIYLSQNPYDCCG
VDGWGALQHGQTVA DWAMVTCNLSSKIIRVTELPGGVPRDCKWERLDLGLLYLVLI LPSC
LTL LVACTVIVLTFKKPLLQVIKSRCHWSSVY

Important features of the protein:

Signal peptide:

Amino acids 1-18

Transmembrane domain:

Amino acids 651-672

N-glycosylation sites:

Amino acids 21-25;74-78;155-159;232-236;292-296;309-313;
312-316;408-414;427-431;500-504;622-626

Glycosaminoglycan attachment site:

Amino acids 533-537

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 566-570

N-myristoylation sites:

Amino acids 24-30;39-45;45-51;141-147;199-205;245-251;
308-314;396-402;416-422;420-426;471-477;
484-490;497-503;522-528;545-551;555-561;610-616

Prokaryotic membrane lipoprotein lipid attachment site:

Amino acids 657-668

Leucine zipper patterns:

Amino acids 48-70;492-514

FIGURE 561

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGCGGTGGCCACAAC**AT**
GGCTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTCTGCTGCTGGGGGCGCTCTGGTGGGTCCCGGG
CCAGTCGGATCTCAGCCACGGACGGCGTTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTG
CAGCATGTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGT
GAATTTTAAAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGGATCCCTTGAAC**T**
TTGGGCTGGAAGTGTTGAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTA**CTTCA**
TAAATACACGGAAGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGG
AGGAAGAGATGATTTTAAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGA
CTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGA
GGAGTCTCGGGGGCGTGA**ACTTGACCCTGTGCCTGAGCCCGAGGCATT**CAGAGCTGATT**CAGA**
GGATGGAGAAGGTGCTTTCTCAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGA
GAGCCACCCTCACACCAGCGGTCC**TGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACAC**
TTTTGAAGAAATTCTGCACGATAAAATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAG
TTCTCCTGCCTCGGTGGAGCGGGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAG
TCAGAGAGGAAGTGGACAGTGC**GTTATT**CATTACAGCAAAGGATTT**CGTTGGCATCAA**ATCT
AAGTTTGTTTTACAAAGATTGTTTT**TAGT**ACTAAGCTGCCTTGGCAGTTTGCATTTTTTGAGCC
AAACAAAAATATATTATTTTCCCTTCTAAGTAAAAAAAAAAAAAAAAAAAAA

567/615

FIGURE 562

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131639
><subunit 1 of 1, 303 aa, 1 stop
><MW: 33900, pI: 4.81, NX(S/T): 2
MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPD
CRFVNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLH KYTEELHIPADET
DFVCFEGGRDDFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSP EESRGRELD PVP
EPEAFRADSEEDGEAFSESTEGLQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDK
LKVPGSESRTGNSSPASVEREKTDAYKVLKTEMSQRGSGQCVIHYSKGF RWHQNL SLFYK
DCF

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-glycosylation site:

Amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 30-34

Tyrosine kinase phosphorylation site:

Amino acids 67-76

N-myristoylation sites:

Amino acids 205-211;225-231;277-283

Amidation site:

Amino acids 28-32

205100 5555001

568/615

FIGURE 563

GCCAGCCGTGGGATTAGGCTTCGCCGGCTACGATTGCGGCCCCCATCTTCTGACTTTTCCTCG
TGTGACCCATCTTTTCAAATTCCCTTACCTGAGGAAGGAGCCCGATTACAAGGATATTTACCT
GCTCCTACCCGTGATCTAGGGACGAGGATGGGAAGACCGCCTGTGGCCATGAGCCCTCCCCGGT
GCTCCTGGGGCTAAGGCTGGGGCTGCAGCCATGGGGGCTGGGTGAGCCCCAGGCCTGGTTGCTG
GGTCTGCCCCACAGCTGTGGTCTATGGCTCCCTGGCTCTCTTCACCACCATCCTGCACAATGTC
TTCCTGCTCTACTATGTGGACACCTTTGTCTCAGTGTACAAGATCAACAAAATGGCCTTCTGG
GTCGGAGAGACAGTGTTTTCTCCTCTGGAACAGCCTCAATGACCCTCTCTTCGGTTGGCTCAGT
GACCGGCAGTTCCTCAGCTCCAGCCCCGCTGTGTGGAGAGGAGCTGCTTGTGGGCAGTGAG
GAGGCGGACAGCATCACCTTGGGCCGGTATCTCCGGCAGCTGGCACGCCATCGGAACTTCCTG
TGGTTCGTGAGCATGGACCTGGTGCAGGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTCGG
GACGCCAAGGTGGAAAGACCGCTTGAGCCCAGGAGTTCGAGGCTGCAATTGAGTTATGATTGCA
CCACTGCACTCCAGCCTGGGCGGCAGAGAAAGGCTCCATCTCTAAAAAAGAAGAGCTAAGTG
CTGTACCTAAAACATGCAGTATATAAACTGGCTGAACTTAGAAATAAACTGTTTTCATGTTAT
GAAAA

568/615

569/615

FIGURE 564

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131649
><subunit 1 of 1, 153 aa, 1 stop
><MW: 17603, pI: 7.42, NX(S/T): 0
MGLGQPQAWLLGLPTAVVYGSLALFTTILHNVFLYYVDTFVSVYKINKMAFWVGETVFL
LWNSLNDPLFGWLSDRQFLSSQPRLCGEELLVGSEEADSITLGRYLRQLARHRNFLWFVS
MDLVQVQWLTPVIPALRDAKVERPLEPRSSRLQ
```

Important features of the protein:

Signal peptide:

Amino acids 1-20

N-myristoylation sites:

Amino acids 4-10;12-18;93-99

Leucine zipper pattern:

Amino acids 102-124

20251004 09:00:00

570/615

FIGURE 565

CGGCACGAGTAAAATGGAGATAATATCACCATGCACTCAGCCCTAGCCACTGCATTGCTGTTA
CTGATACCATTACTGCTGCTACGTCGTTTTTTTGATGGCTCAGCCCTTAGGGAAGGGGGATCA
AGGGAGAAGCCCGGACCTTCCCGCAGGAGGTGGGCTGGGCACAGCCCTGAACCATGGAGGTCA
CCCACCCTGAGGTCGGGACCTGGGTTCCCTTCCTATCCACTGGGGGTCCCAGCCTTTGTCTTC
ATCTCTCCAGGTCCCAGCCCTTCACAGTGGGCACTTCCCTGCCTGTGACGGAGGCCCCAGCCA
TCTCC

CGTCGAGTCTG

571/615

FIGURE 566

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131652
><subunit 1 of 1, 89 aa, 1 stop
><MW: 9688, pI: 11.49, NX(S/T): 0
MHSALATALLLLIPLLLRRFFDGSALREGGSREKPGPSRRRWAGHSPEPWRSPTLRSGP
GFPSYPLGVPAFVFISPGPSPSQWALPCL
```

Important features of the protein:

Signal peptide:

Amino acids 1-18

Glycosaminoglycan attachment site:

Amino acids 58-62

2005-04-01 10:00:00

[illegible]

FIGURE 567

AGTCTAGCAGGAAAGGAGAGGGAGCTTTTCCCCGAAGACCCTCCTGGACCAGCCCCAGGCTCCT
GTGCTGGTTGCACGCCAGGGCCTGTACTGACCACCTCCACGTGCCACTGGGGCTGTAAGGAGGA
ATGGCGGCCGTGGGCAGCCTGCTTGGCCTGGCAGCCTCTTCTGGCTAGGGGGGCCAGAACGCC
TCTGACCACAGCCTGTGGCTCCTGAGGAAGCCCCGAGGCTCATCCTGCCCCGGCACGGGTAC
CAGCTCTGCCGGCTGAGGCAGAGCACCGTGAAGGCCACCGGACCTGCACTCCGCCGCTGCAC
ACATCCTCCTGGCGAGCTGACAGCAGCAGGGCCTCACTCACTCGTGTGCACCGCCAGGCTTAT
GCACGACTCTACCCCGTGCTGCTGGTGAAGCAGGATGGCTCCACCATCCACATCCGCTACAGG
GAGCCACGGCGCATGCTGGCGATGCCCATAGATCTGGACACCCTGTCTCCTGAGGAGCGCCGG
GCCAGGCTGCGGAAGCGTGAGGCTCAGCTCCAGTCGAGGAAGGAGTACGAGCAGGAGCTCAGT
GATGACTTGCATGTGGAGCGCTACCGACAGTTCTGGACCAGGACCAAGAAG**TGA**CCGTGGCTC
CAGCCACCCCGGGACATTGCTAAGATGGGAGGGCTGTTCTTAAATCACTCGTTCTTGAAGCTGC

573/615

FIGURE 568

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131658
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18903, pI: 11.08, NX(S/T): 1
MAAVGSLLGLAASSWLGGQNASDHSLWLLRKPRGSSCPGTGHQLCRLRQSTVKATGPALR
RLHTSSWRADSSRASLTRVHRQAYARLYPVLLVKQDGSTIHIRYREPRRMLAMPIDLDL
SPEERRARLRKREAQLQSRKEYEQELSDDLHVERYRQFWTRTKK
```

Important features of the protein:

Signal peptide:

Amino acids 1-18

N-glycosylation site:

Amino acids 20-24

N-myristoylation sites:

Amino acids 5-11;9-15;17-23;18-24

20570-01503

FIGURE 569

GGTGCCAAGGGTTTCGGGGGGGAGCACTGAGGCTTTAGCAGCTCTCCTGTATCCTCATTGTCAT
 CCTCCTGTAGCAGCTGGAAAAATTCAGATTACAGGTGAAATTCCTGGCTGGCAATCTTCTGTA
 TATGGACACAGTGATGTGCCAGAAGGGCTTTGCATCCCTGAGACTGAAGGAAGCTCCATTTTT
 GGAGCCCTCCCACACCTTGCTCTGTGTGCCTCTCATTCTGATTTGAATTCTTATTTTGGCTATA
 TGATGAAGCTGTAATCCTAAGTTTAAAAAGGGGAGTAGGTATTGACATCATGGTAGAAATAGG
 CTGTCTTATGGAAGTGTAGTTAGGGATCACAGCCTATTGGACCAGCCCCAGCCTTAGCAGCAG
 TTCTGTACACTGATTCTTCCAGATTAGTCTACGTTCCCTCGAACAGACCTATGCCATGGGTTA
 CAACTACAATTTGTTGTGCGATTAGAGTTAACTTACAGACTCTCAAAACCCCATTTCTTTGGGTT
 TAGGCAACTTCCAGAAAGTAGTCATTTATTTGAATTTTAGTCTAAGATCAACTGAATTAGGGAG
 GTTTGAAAGTGTAAGCAAAATCGTACATTTCCAAACACTTTGTAAAGAAGGAATGGGTTAGTG
 TCAACTAAAGGAAATGGTGTGCATCCAGCAAAAGAAAGAGACCGAAAGCAAAGTCATAAACC
ATGCCACAGAGCTCAGCTGTCCTGCTCCGTGTCCTCTCCATACCCTTGTTGACTGTGCTCATA
 TTAGCCAGAGACCTAAGTGCTCTTGGAGGATGTCCCTGGGGCCCCCTCCCCCTCCGCTGTCAC
 TGTCTACTTCCTGATCCTCTCTTCTGTGCAGGAGAGGTCCAGGCCTTCTATGAGGACCTGAGT
 GGCCGGCAGTACGTGAATGAAGTCTTCAACTTCAGCGTGGACAAGCTCTATGACCTCCTCTTC
 ACCAACTCGCCCTTCCAGCGGGATTTTCATGGAGCAGCGGCGCTTCTCTGATATCATCTTCCAT
 CCATGGAAAAAGGAGGAGAATGGAAACCAGAGCCGAGTGATTCTTTACACCATCACCTTACC
 AACCTCTGGCTCCCAAACTGCCACTGTGAGGGAGACACAGACCATGTACAAGGCGAGCCAG
 GAGAGTGAATGTTACGTGATAGATGCCGAAGTCCTCACCCACGACGTGCCCTACCACGACTAC
 TTCTACACAATCAATCGCTACACGCTCACCCGTGTGGCTCGGAACAAGAGCCGACTCAGGGTC
 TCCACAGAGCTGCGCTATCGAAAACAGCCCTGGGGGTTAGTGAAAACGTTTCATCGAGAAGAAC
 TTCTGGAGTGGGCTGGAGGACTACTTCCGCCATTTAGAGAGCGAGCTGGCCAAAACGGAGAGC
 ACTTATTTGGCTGAGATGCACAGACAATCTCCAAAGAGAAGGCCAGCAAGACTACAACGGTG
 CGGAGGAGGAAGCGTCCCCATGCCACCTGCGAGTCCCTCACCTGGAAGAGGTGATGAGCCCG
 GTCACCACGCCACAGATGAGGATGTGGGCCACAGGATCAAACATGTGGCAGGTTCCACACAG
 ACGCGGCATATCCCGGAGGACACCCCCAACGGTTTCCACCTGCAGAGCGTGTCCAAGCTGCTG
 CTGGTTATCAGCTGTGTTCTGGTGCTGCTGGTCATCCTTAACATGATGCTCTTCTACAACTC
 TGGATGTTGGAATACACCACGCAGACCCTCACTGCCTGGCAGGGTCTAAGGCTCCAAGAAAGG
 TTACCCCAGTCTCAGACAGAATGGGCCCAGCTCTTAGAGTCCCAACAAAAGTACCACGATACT
 GAGCTCCAAAAATGGAGGGAAATCATCAAATCCTCAGTGATGCTCCTTGACCAGATGAAGGAC
 TCGCTCATCAACCTTCAGAACGGCATCAGGTCCCGCGACTACACGTGCGAAAGTGAAGAAAAG
 AGGAATCGCTATCATTGACAAGGCAGGAACAGGGTGGCTGCAAGAGGCCTGTGCAATACATGT
 ACATAGACCATATAAATATATATATATAAATATATATATATACAGAATATAAATATATATATT
 ATATACAGATTTTAAAAAAGAGATAATGCCTATGTACCAGGGAGAAGGAGCGGGCCCTCCCGC
 GCCCTGTGCTGGCCGAGCAGCGTTTTCTTATGGTGGAGCAGCTGAGGAGGGCAGGAACCGCC
 TCTCAGCACCGACCTCCCCTGATCTCCCTCCTCCACCTCTGTTCCCCACCCCTTCCCTTGC
 TGGCCATTCTTGGCTTTTAGAAGGGAAATGTTGAGCCAAAGTTATGCCTGCGAAGACCCTAAG
 GTCTCAAAAAGAAGTCTTAAGACGGCATTGCTTAAGGTGCTTCATTCCCTAATCCCCTTTTGA
 TTTGTTTCCAAAATAAAAGAGAATCTTTTCTCCCTAAAAA

575/615

FIGURE 570

><subunit 1 of 1, 425 aa, 1 stop
><MW: 49786, pI: 8.84, NX(S/T): 3
MPTSSAVLLRVLSIPLLTVLILARDLSALGGCPWGPLPLRCHCLLPDPLFCAGEVQAFYE
DLSGRQYVNEVFNFSDKLYDLLFTNSPFQDRDFMEQRRFSDIIFHPWKKEENGNSQSRVIL
YTITLTNPLAPKTATVRETQTMKASQSECECYVIDAEVLTHDVPYHDYFYTINRYTLTRV
ARNKSRLRVSTELRYRKQPWGLVKTFIEKNFWSGLEDFRHLSEELAKTESTYLAEMHRQ
SPKEKASKTTTIVRRRKRPVPHLEEVMSPTTPTDEDVGHRIKHVAGSTQTRHIPE
DTPNGFHLQSVSKLLLVISCVLVLLVILNMMMLFYKLWMLEYTTQTLTAWQGLRLQERLPQ
SQTEWAQLLESQQKYHDTLQKWREIIKSSVMLLDQMKDSLINLQNGIRSRDYTSESEEK
RNRYH

Important features of the protein:

Signal peptide:

Amino acids 1-28

Transmembrane domain:

Amino acids 312-334

N-glycosylation sites:

Amino acids 73-77;114-118;183-187

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 97-101

Tyrosine kinase phosphorylation sites:

Amino acids 144-153;188-196

N-myristoylation sites:

Amino acids 201-207;291-297

Leucine zipper pattern:

325-347

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

576/615

FIGURE 571

GTAGAGAGTGAAGCAGCAAGACTGCAGAGCCTCATCAAGAAGTGTGGAGTGAAGGGAAGGCTTCAGATGGACAAT
TTGTGTGCTGGGGAAAAAATGGAATGTGCTGCAAATCCCCCTGTGGATAAGGGTGGACGGCTGCTCTGTCAACTT
TGACCATTTTCAGATTCTGCGGGCCATTGGTAAAGGGAGTTTTGGAAAGGTATGCATCGTCAGAAAGCGAGACAC
TAAGAAAAATGTATGCAATGAAGTACATGAACAAGCAGAAGTGCATCGAGAGGGATGAGGTTTCGGAATGTTTTCCG
GGAGCTGCAGATCATGCAAGGGCTGGAGCACCCCTTCCTGGTCAATCTGTGGTACTCCTTCAGGATGAGGAGGA
CATGTTTCATGGTGGTGGACCTGCTCCTGGGAGGCGACCTGCGCTACCATCTGCAGCAGAATGTGCATTTTCACAGA
GGGACTGTGAACTCTACATCTGTGAGCTGGCACTGGCCCTGGAGTATCTTCAGAGGTACCACATCATCCACAG
AGACATCAAGCCAGACAATATCCTGCTGGATGAACACGGACATGTTTACATTACAGACTTCAACATAGCGACGGT
AGTGAAAGGAGCAGAAAGGGCTTCCTCCATGGCTGGCACCAAGCCCTACATGGCTCCAGAAGTATTCAGGTGTA
CATGGACAGAGGCCCGGATACTCGTACCCTGTGCACTGGTGGTCCCTGGGCATCACAGCCTATGAGCTGCTGCG
GGGCTGGAGGCCGTACGAAATCCACTCGGTACGCCCATCGATGAAATCCTTAACATGTTCAAGGTGGAGCGTGT
CCACTACTCCTCCACGTGGTGAAGGGGATGGTGGCCCTGCTGAGGAAGCTCCTGACCAAGGATCCTGAGAGCCG
CGTGTCCAGCCTTCATGACATACAGAGCGTGCCCTACTTGGCCGACATGAACCTGGGACGCGGTGTTCAAGAAGGC
ACTGATGCCCCGCTTTGTGCCAATAAAGGGAGGTTGAACTGCGATCCACATTTGAGCTTGAAGAGATGATTTCT
AGAATCCAAGCCACTTCACAAAAAGAAGAAGCGATTGGCAAAGAACAGATCCAGGGATGGCACAAGGACAGCTG
CCCGCTGAATGGACACCTGCAGCACTGTTTGGAGACTGTCCGGGAGGAATTCATCATATTCAACAGAGAGAAGCT
CAGGAGGCAGCAGGGACAGGGCAGCCAGCTCTTGGACACCGACAGCCGAGGGGGAGGCCAGGCCCAAAGCAAGCT
CCAGGACGGGTGCAACAACAACCTCCTCACCCACACCTGCACCCGTGGCTGCAGCAGCTGAGCCCCACACTTGTTG
CTGCTCAACAGGACTGCACTCGTCTCTGCCCTGCCACCCAGAGCCCCCTCTTTGTGCCCTGATGGTCCCTGTCTC
ACCCCTGAAAAACATCAGATGCAGAAAAAGCCCTGGACTTGGAGCTGGGAAGCCTGGGTCTGTGCTCCCATCTCCAT
GACTGATTCACGTGTGACCTCAGACAAGTCAGCCCTCTCTGTGCCTCCGTTTTCTGCATCTGCCAAAAGGGGTTA
AACACTTCTGCCCACTTCAAATTACAAGATTATGGGGAGAACCCCAATTAGGTAGGAAACATGAAAAACCTTTGA
TATTTATAAAATCATTTTTACGTGCAAAATATAACCTTAATATTTGAAGTGACCCCCATTCCCCAAAGCAATCAA
ACCGTCATGACTTTTGCAATTTGGCACATCCTAGCTTGTTAGAGGGCACCTCCGAAAAACACAGCCCTGACAGCAA
AATAAAGGTCTGATATGTTGGCCCCCTTCTATGGAAACAACGCTGCCAAATCCTGGAGCAAAACCTGAAGTGTCTT
CATGTGCATTTCTCTGGCAGGCCACAGTCTTCTGAGCTTGTAAGATGGTGCAGCATGCAGACCAGACTTGTCCCC
AAGGTCTCAGCGCTGCGGTCTCACTCCTCCCTCATTTAAGAAAGACTATCCTTACCTTTTAGTTTCAGCAGTCCCT
CACCACCACCATATCCCCAGTGCTGGGATGGCACACAGGTGTCCATTGAGATGAGAGTTGGGTGCTGAGCATTG
GTTACTCCTGCAGAGTGTAATCAGCACCCCATCCAAGTGGCCCGAAAGCCAGACCTGCAGCAGAACTCTCCAAC
TCTCTATCAGCTTTCAGGGTTTTCTCTCCTGGGAAGGGTGTAAATCAGCTTGTGATTTCTTCTTACAGAGAGT
ATCCAATCGGTATTGGTGGAGCGGCTCCCTATTTATACAATAGGAAGCATGGGTGCTTAGAAAGTTTATTTCAGG
AGGAAAATGGGTTTACACAAAAAGCAAACCTACATTTCTGATCTGCTCAGGGAGAAGCTTGCTTTGAACTGGGAAGA
TGTTGGGATGAGCAGGGAAAGCTTAGACTTTGGAGTCAGGTTTGTGTTTCAAGATCCAGCCCTGCTGGCTACTAAC
TAACCTGGGAGACCTTAGGCAAAGCATGCAATCGCTCTGAATGGCAGTTTCTCATTTTTTAAACAGGGATAATAAA
ACTAATATTGCAGGGGAGTTACAGGGTTAAATAAGATCCTGTGTGTAACCCCAAGCATTGGATGACTCATAGAAT
GGCCTTTTTTGTGAGCATAATCGTCATCATATTATTTAGATACTTTCTTCTTCACTCACCCAGCAGGTGAGTTTTT
TGTGCAACAAACCTGTTTGGATTCTTCCAAATGTTCTTCTGGGGTCTTTGATATTTGTTTGTACATCCTGC
TGAAGTTCGACTGTGTTTTTATTTTTTATCCAACCTTCCATTTTTTCACTTTTTTACATGATTACTCAATCCTTGGG
GCTGTCCATGTGATCTCTTAGATTTCTTAAAGACATTTTAAATGTATGGTTAGGTTTTATATTTTATTTTTTAA
AAAAGAAATAGTCAGTGTTTTCTCCTTTCAACCGAGACTATTTCTGGATTGTGTGCTCCTCGTCAGTTGACTTGT
TTTGACACTTTTTCTTTACTTCATGTCCCCATCAACAACCGTCTCTGCTCCCCACCTCCCCAGGAAATAAGGGGC
CTGCTCCTCTCCCTACTGTGACCTGGAGGCTCTTAAGATGATGATGTTTTTTTTTATTGGGCTGAGTTTACAGAA
TTAGGGGCAGGAGCTGGAAGTCGCCCTAGGAACACCAGATTTTCTGGTTCTGTTTCAAGTTGGCATTTCCTGTTTG
GAATAAACTATTTCTTGG

577/615

FIGURE 572

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA136110
<subunit 1 of 1, 364 aa, 1 stop
<MW: 42195, pI: 7.40, NX(S/T): 1
MKYMNKQKCIERDEVNRNVFRELQIMQGLEHPFLVNLWYSFQDEEDMFMVVDLLLGGDLRY
HLQQNVHFTTEGTVKLYICELALALEYLQRYHIIHRDIKPDNILLDEHGHVHITDFNIATV
VKGAERASSMAGTKPYMAPEVFQVYMDRGPYSPVDWWSLGITAYELLRGWRPYEIHVS
TPIDEILNMFKVERVHYSSTWCKGMVALLRKLLTKDPESRVSSLHDIQSVPYLADMNWD
VFKKALMPGFVPNKGRLNCDPTFELEEMILESKPLHKKKKRLAKNRSRDGTDSCPLNGH
LQHCLETVREEFIIFNREKLRRQQGQGSQQLDTSRGGGQAQSKLQDGCNNNLLTHTCTR
GCSS

Important features of the protein:

N-glycosylation site:

Amino acids 285-289

N-myristoylation sites:

Amino acids 123-129;290-296;337-343;339-345;348-354

Serine/Threonine protein kinases active-site signature:

Amino acids 92-105

FIGURE 573

CTCCAGTTTCGCCGACTGTAACATGTTTTCATCCAGTTTCAGTATGTTTTGTATGCAAGTTGGAAATAAAATAAACGTC
 CTGAACTGGATGAAACATGTTACAGTCGGCCGAAACATGAGAGGCTGTGTGAGAAGCTGCAGCCGCCCGCAGAGG
 AGACCTCAGCATCATCTAGAGCCCAGCGCTGGCCCTGCCTCCGCCTGCGCCGCCGCCCGCGTCCGCGTTTCTGTT
 CCTGCTACTGTCCCACCTAAACAACCTCCCGTTACACGGACAAAGTGAACATCTGTGGCTGTCTCTCTTTTCTTC
 CTCTCTTTCCAACCTCTTCTCTCTCTCCACTTCCAGCCGCAGCAGAAAAGCCCCAACCCAACTGACGCTGGCA
 CAACTGCAAACGGTGTCTCCGCACAACCTTTATCTCGCTCCTCGGGCTCCCTTAAGGCATTGGACCCATCGCCGC
 GTCTTTTATTTTTGCAAAGTTGCATCGCTGTACATATTTTTGTCCCCGCCACCTCCCTCTGTCTCTGGAGTGGCC
 TACAGCCCCGCAAACCTCTCTTGGAGCTGCGCCCTAGTGCCCTGTCTGGGCAGTGGCGTTCCCCCCCATCTCTCC
 GCGCCAGCCCCCTGCTGCTCTGGGCAGACGATGCTGAAGATGCTCTCTCTTTAAGCTGCTGCTGCTGGCCGTGGCT
 CTGGGCTTCTTTGAAGGAGATGCTAAGTTTGGGGAAAAGAAAAGGAGCGGAGCAAGGAGGAGAAGGTGCCTG
 AATGGGAACCCCCCGAAGCGCTGAAAAGGAGAGACAGGAGGATGATGTCCAGCTGGAGCTGCTGAGTGGGGGA
 GAGATGCTGTGCGGTGGCTTCTACCTCGGCTGTCTGTGCTGCGGAGTGACAGCCCGGGGTAGGGCGCCTG
 GAGAATAAGATATTTTTCTGTTACCAACAACACAGAATGTGGGAAGTTACTGGAGGAAATCAAATGTGCACTTTGC
 TCTCCACATTCTCAAAGCCTGTTCCACTCACCTGAGAGAGAAGTCTTGGAAAGAGACCTAGTACTTCTCTGCTC
 TGCAAAGACTATTGCAAAGAATTCTTTTACACTTGCCGAGGCCATATTCCAGGTTTCCCTTCAAACAACCTGCGGAT
 GAGTTTTGCTTTTACTATGCAAGAAAAGATGGTGGGTTGTGCTTTCCAGATTTTCCAAGAAAAACAAGTCAGAGGA
 CCAGCATCTAACTACTTGGACCAGATGGAAGAATATGACAAAGTGGAAAGAGATCAGCAGAAAAGCACAAAACACAAC
 TGCTTCTGTATTTCAGGAGGTTGTGAGTGGGCTGCGGCAGCCCGTTGGTGCCCTGCATAGTGGGGATGGCTCGCAA
 CGTCTCTTCATTCTGGAAAAAGAAGGTTATGTGAAGATACTTACCCCTGAAGGAGAAATTTTCAAGGAGCCTTAT
 TTGGACATTACAAAACCTTGTTCAAAGTGAATAAAGGGAGGAGATGAAAGAGGACTGCTAAGCCTCGCATTTCCAT
 CCCAATTACAAGAAAAATGGAAAGTTGTATGTGTCTTATACCACCAACCAAGAACGGTGGGCTATCGGGCCTCAT
 GACCACATTCTTAGGGTTGTGGAATACACAGTATCCAGAAAAAATCCACACCAAGTTGATTTGAGAACAGCCAGA
 GTCTTTCTTGAAGTTGCAGAACTCCACAGAAAGCATCTGGGAGGACAACCTGCTCTTTGGCCCTGACGGCTTTTGT
 TACATCATTTCTTGGTGATGGGATGATTACACTGGATGATATGGAAGAAATGGATGGGTTAAGTGATTTTACAGGC
 TCAGTGCTACGGCTGGATGTGGACACAGACATGTGCAACGTGCCTTATTCATACCAAGGAGCAACCCACACTTC
 AACAGCACCAACCAGCCCCCGAAGTGTGTGCTCATGGGCTCCAGCATCCAGGCAGATGTGCTGTGGATAGACAT
 CCCAATTGATATAAACATCAATTTAACGATACTGTGTTTCAGACTCCAATGGAAAAAACAGATCATCAGCCAGAATT
 CTACAGATAATAAAGGGGAAAGATTATGAAAGTGAGCCATCACTTTTAGAATTCAAGCCATTTCAGTAATGGTCCT
 TTGGTTGGTGGATTTGTATACCGGGGCTGCCAGTCAGAAAGATTGTATGGAAGCTACGTGTTTGGAGATCGTAAT
 GGGAAATTTCTTAACCTCTCCAGCAAAGTCTGTGACAAAGCAGTGGCAAGAAAAACCACTCTGTCTCGGCCTAGT
 GGGTCTGTAGAGGCTACTTTTCCGGTTCACATCTTGGGATTTGGAGAAAGATGAACTAGGTGAAGTTTACATTTTA
 TCAAGCAGTAAAAGTATGACCCAGACTCAAAATGGAAAACTCTACAAAATTTGTAGATCCAAAAGACCTTTAATG
 CCTGAGGAATGCAGAGCCACGGTACAACCTGCACAGACACTGACTTCAGAGTGCTCCAGGCTCTGTGAAACGGC
 TACTGCACCCCCACGGGAAAGTGCTGCTGCAGTCCAGGCTGGGAGGGGGACTTCTGCAGAACTGCAAAATGTGAG
 CCAGCATGTGCTCATGGAGGTGTCTGTGTTAGACCGAACAAGTGCCCTCTGTAAAAAAGGATATCTTGGTCTCAA
 TGTGAACAAGTGGACAGAAACATCCGCAGAGTGACCAAGGCGAGGTATTCTTGATCAGATCATTGACATGACATCT
 TACTTGCTGGATCTAACAAGTTACATTGTATAGTTTTCTGGGACTGTTTGAATATTTCTATTCCAATGGGCATTTAT
 TTTTTATCTGTCTATTAAAAAAGACTGTTATCCTGCTACACACTCCTGTGATTTTATTCTCTTTTATTAA
 TTTAAAAATAATTTCCAGAAATGTGCAGATCCTCTGTGTGTATGTGCAGCATGTTTGTTCACATATGCACATACAC
 ATACTCATAACCCCTATATGCGTTGTTGCATAACAGATGATTTTTTAAATATATACTTCTTATGCAAAGTAAT
 TTACACAGAAATTCATTGTAAATGATAATGGATTTTTTATGTTACTAGAAAGAGATTATTTGACTTCCCAGGAA
 TTTTCTGTCTGTAATCACTAAAGTCAACTTTAATAGAGTTTTGAAACAGTACTGTGCAATCCGATGGATCTAATT
 AAAAAAAGGCAATATTTTTATATTAAAGTACTATACTAGGAGAGAATGTTTCAGAACTCCCTGATGAATTTCTA
 AGTGAGCAACTTGATATAAAATTTGTAATCTTCATTTTTGTGTCAGTGTATCCAGTTACAGAATGCTACACACTTACC
 TTTTATTGGCTGAGAAATCTGGTTATTTTCATCTTAATCTCAAGATTGTTTTCAAGTGTTTTATAATTAAATCAT
 AATAGCATATTTTAAAAATCAAAAA

579/615

FIGURE 574

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139592
><subunit 1 of 1, 882 aa, 1 stop
><MW: 98428, pI: 8.89, NX(S/T): 5
MKHVTVGRNMRCVRSRQRRPQHHLEPSAGPASACAAA VAVSVPATVPPKQLPLHG
QVNICGCPLLFFLLFQLLLLLLPLPSRSRKPTQLTLAQLQTVSSAQLYLAPRAPLRHWT
RRVIFYCKVASLYIFLSPPPPSVSGVPYSPANSSWSCALVPLLGSVPPHPPAPSPCCSG
QTMLKMLSFKLLLLLAVALGFFEGDAKFGERNEGSGARRRRCLNGNPPKRLKRRDRRMSQ
LELLSGGEMLCGGFYPRLSCLLRS DSPGLGRLENKIFSVTNNTTECGKLLEEIKCALCSPH
SQSLFHSPEREVLERDLVLP LLCKDYCKEFFYTTCRGHIPGFLQTTADEFCFY YARKDGGL
CFPDFPRKQVRGPASNYLDQMEEYDKVEEISRKHKHNCFIQEVVSGLRQPVGALHSGDG
SQRLFILEKEGYVKILTPEGEIFKEPYLDIHKLVQSGIKGGDERGLLSLAFHPNYKKNGK
LYVSYTTNQERWAIGPHDHILRVVEYTVSRKNPHQVDLRTARVFLEVAELHRKHLGGQLL
FGPDGFLYIILGDGMITLDDMEEMDGLSDFTGSVLRDLVD TDMCNVPYSIPRSNPHFNST
NQPPEVFAGHLHDPGRCAVDRHPTDININLTILCSDSNGKNRSSARILQIIKGKDYSE
SLLEFKPFSNGPLVGGFVYRGCQSERLYGSYVFGDRNGNFLT LQQSPVTKQWQEKPLCLG
TSGSCRGYFSGHILGFGEDELGEVYILSSSKSMTQTHNGKLYKIVDPKRPLMPEECRATV
QPAQTLTSECSRLCRNGYCTPTGKCCSPGWEGDFCRTAKCEPACRHGGVCVRPNKCLCK
KGYLGPQCEQVDRNIRRVTTRAGILDQIIDMTSYLLDLTSYIV

Important features of the protein:

Transmembrane domains:

Amino acids 63-80;186-201

N-glycosylation sites:

Amino acids 152-156;281-285;598-602;629-633;641-645

Glycosaminoglycan attachment site:

Amino acids 417-421

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 856-860

N-myristoylation sites:

Amino acids 12-18;413-419;457-463;698-695;720-726;723-729

EGF-like domain cysteine pattern signatures:

Amino acids 805-817;837-849

FIGURE 575

CGGCTCGAGAGCGGGGCAAACCTGCTTGGCACCTCTTCAATAGGTGACATTCAATGATAGATCT
 CTGGCTTCCCTGCTCTGTTTGTCTGGTTGCCCTGGAAAGCCTGCTGCTCAGCCCATGCCCCGG
 GACTTCCTCCACCCTCACCAGGACATTCTTTCCATCTCTTGTCTCCTGTGTGCAAGTCCCTTT
 CTCCTGGATTCCATGTCTTGAATGTTTCTTAATTTACTTCCTCATTTTGGCAGAGGATGTCCT
 CCAGTTGTTTTCTGGGAATGCTAATATGCAAGTGAACCAGTGAACCTGCAGTTCTGCCCCACACA
 GGGTTAATAACCAATCAGATTCTCTCTTTTCAAGATGGTTAACATAACAGACACCAAGAAAGG
 GAAGAGGAGCCGACAGCAGAGGGGGAAGCTGAAAAGACGCACAAAGAATGGCCATAAAAGATA
 TGAGCAACCCCAGCTTTCCAGACAGTCACTTTTCCCAGTGGTCATACCTGGTCTGGAAGATTC
 CCCATCATCTCGAATAAAGCTGTTGTTGCTTTTAACTCCATGGAGAGACCGAATGGAGTGAGC
 CCAGCAGGGCATGCTGGGCAAGAGAGGTCCCCGAGTCCCAAATAAGAATTTCAACTAGTATA
 AAACGAGGCAGCGAACCCACACGTGGAAGTCTGATACCGCTTGCAGAAGGGAATTGAATAGAT
 GTCTCCCTATTGGTAAGGATGTGGTTTTATTGACTTGAAATAACAAAGCCCCGCAAGCAACAAC
 TGATCATCCGCGGGATGCTGCCACAAGGAATAATTGAGCACTCATTTCAGACACAGGGGAAACC
 ACTGCCTCTTTTCACTCTTTCTCCCAGATTCCAACAGTCAGTGTTACAGCATTTTACCTTGTTT
 ACCTCCCTGAGAAGACGTTGCAGCTCACTCACCCAGTGGGCACTGGGAGCCTCTGCTCAGGT
 GGGAGACAGATGCCCCCACATGCACATCTGGTGTATTATGAAGCAGATACTGGGGCTTCATAA
 ACACAGAAGGGGCAGGGAAGTAGCCCCAGGGCATAGTGTGGGGCCTCCTGACTAAAAGTAGCTT
 GCAAACCCCTGCCTATAACAGCCACTTCCTGGCAGTCATTGTGCCACTTAGGAGCCCTCCTCA
 CCGCTCTTCTTTAGTTCTTCACTCTGTATAGACCCTGCCAGAGCAGCTCAGGGTGGAGCAG
 CTGCAGCCATGGGACCTGCTCCAGGCAAGGCCCTATGCTACACAGTCCCTGGGGGTGAGGATT
 CCAGAGGAGTCAATTCCTTACCCGCCCTTCCAGGGGATGGCCGAGATGAGCGTTCCACAGGG
 AAAGTGAAAGTTGCAAGGCTGTTGCGGAAACACCAGGGGTTCATTTAGGTCTTGCTGCTCAT
 CTCACAGAAAGCCAGTCACTGAGACAAGTATTGTGAGGGAAGAAGGCTTTATCCAGGTGCTAC
 AGACAGGGTGAACAAGAGATCAGTCTCAAATCCATCATCATCAACTGACTAAAATTAGGGGTT
 TATATAGCAGGGAAGAAATGTAACCTACATGTGGGAAAACAGGGATTAACGAGGGGCAAGGAAG
 AGGAGTTGGCCAACAGGCAGCAGGTGGTCACCTGGGGAATCATGATGGGTGAGGGGTCTGGCT
 TCTCACTTTCAGATGTGGGGATCTGGTAAATTTTCACTTTCTTGATACCGTCTGGGAGGATTG
 CTGGCTGGTTTTCTGAGAAAGGAATTCAGATGACATAAATGTAAATTTCTCCTTGGGTTTTCAA
 GACTGAGAGGGTCAATTTCTAGGTTTATTCAAGAAAAACCATAAACATCAGTTCTATGGGACA
 ATTGGGCCCATTTCAGGCTCTGAGGATAAGGGTTAATGAGGGGACAGAGTCGCCTGGAGAAG
 TTCCTGAGGGCCTACAAGAACTAGAGAGGCTTCCTGGCAAAGCTCTATGCTGTCTATCCTCT
 CTTCTCTCCTTGAGGAAGATTCCAGTATAATAGACCCGAGGTGAAAAGGCTTTTGTTCATAA
 AGTAGAAAACCTGAAGGGGGTGGGAGGCACATGGATTTGAACCAGAGACCGCTTGGCCTGCAG
 GCAAATGCTGTACCTTCAGTTGCACCCCTCACTTGTTACAGCTGTTTCTGATAAGCACTTGTTG
 CAGCCCCATCAGCACCTCGATTTCTTCTTGGTGAGTCCATGGGAACAGCCCCACTGCAAACAA
 CCCATTCTGCTCTCCTCTTTCTTCTAAACCTCAACCTCCTCCTACCTGGCAGTCCACAGGCC
 TACAGCTTCTCCTCAGTGGGAAAGACATCAGCTTGGAAAACCACTTGGAAAGCCAACGTTATC
 CTAGAAAAGCTTTTTTAAATGACCCAGCAGGACAAGTCTCCGGATGGCCTTGGCCAACCCGGTG
 CTTCCCTCTTTTCTTGGTTGTAGTTCTCAGAATAACTAGAGAATGTACTGGGAGTGTTGTCTT
 GAGATAAGGAGGAACCTGTCCTAAACCTGGACTCTGTTCCCATCACACCTAGAACAGGATGTCC
 TGCAACGCTTTAGCCCAATGATCCAAGTTGCCCTTGGGGTATAAACTTGACAGCAGAGGGCG
 TTCAGGGTCCCTCAGCTGCAGTGTGAAGTGGGACACACAGGTGAGACTCCATCTGCCCTGGGC
 AGGTTCTTGAGCCTTGGGGGACCAGTTTACCCTACATCCCAGGCTTCTGTTGTCCCTTGCCCTG
 CCTGTAAGGAATAAAGTTGCTTTGCTTA

581/615

FIGURE 576

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139608
><subunit 1 of 1, 80 aa, 1 stop
><MW: 8927, pI: 3.77, NX(S/T): 0
MIDLWLPALFVLVALESLLLSPCPGTSSTLTRTFFPSLVSCVQVPFSWIPCLECFLIYFL
ILAEDVLQLFSGNANMQVNQ
```

Important features of the protein:

Signal peptide:

Amino acids 1-29

Transmembrane domain:

Amino acids 47-62

N-myristoylation sites:

Amino acids 25-31

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 577

ATCGGTTAGCGCCTTGCCATGATTAATCCAGAGCTGCGGGATGGCAGAGCTGATGGCTTCATA
CATCGGATAGTTCCCAAGTTGATACAAAACCTGGAAGATTGGCCTTATGTGCTTCCTGAGTATT
ATTATTACTACAGTTTGCATTATTATGATAGCCACATGGTCCAAGCATGCTAAACCTGTGGCA
TGTTTCAGGGGACTGGCTTGGAGTGAGAGATAAGTGTTTCTATTTTTCTGATGATACCAGAAAT
TGGACAGCCAGTAAATATTTTGTAGTTTGCAGAAAGCAGAACTTGCTCAGATTGATACACAA
GAAGACATGGAATTTTTGAAGAGGTACGCAGGAACTGATATGCACTGGATTGGACTAAGCAGG
AAACAAGGAGATTCTTGGAATGGACAAATGGCACCACATTCAATGGTTGGCCATCAAACCTCC
AAATGGTCTTGCAACTGGAGCCTCCGACAATGGCTTCTTCTGCTGGGACCCCTTAGATTAGGCC
TCTGAGGGAGCTCTGACTGCCGTTTCCCCAAAACAATGTCCCCTGTCAGCAGGAAGCAGTTAA
ATCAGTCTTCATCCTTATCCTTAATATAACGGCAGTTAGATGTACTTCTTTAGAGGGAGTAAA
TTTATCAATTCAGAGCAATTCATCCTCCTCTTCCATCTTTGATTACAGTTAATAGGCTATA
AATTTTGATAATGTAGAATAAACTACAGAAAACCTTCTTG

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 578

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA143292
><subunit 1 of 1, 160 aa, 1 stop
><MW: 18576, pI: 9.29, NX(S/T): 3
MINPELRDGRADGFIHRIVPKLIQNWKIGLMCFLSIIITTVCIIMIATWSKHAKPVACSG
DWLGVRDKCFYFSDDTRNWTASKIFCSLQKAELAQIDTQEDMEFLKRYAGTDMHWIGLSR
KQGDSWKWTNGTTFNWGPSNSKWSCNWSLRQWLLLLGPLR
```

Important features of the protein:

Signal peptide:

Amino acids 1-42

N-glycosylation sites:

Amino acids 78-82;130-134;146-150

N-myristoylation site:

Amino acids 131-137

FIGURE 579

TGAAGGCCTGTGAGTGAGGAATGCCTCTCACCAGCTGTGCCTGAGCTGCAGCACTCCAGCCAC
TGCTGTCTCCTTAGCTGCTCACATATGGATACTTTCACAGTTCAGGATTCCACTGCAATGAGC
TGGTGGAGGAATAATTTCTGGATCATCTTAGCTGTGGCCATCATTGTTGTCTCTGTGGGCCTG
GGCCTCATCCTGTACTGTGTCTGTAAGTGGCAGCTTAGACGAGGCAAGAAATGGGAAATTGCC
AAGCCCCTGAAACACAAGCAAGTAGATGAAGAAAAGATGTATGAGAATGTTCTTAATGAGTCG
CCAGTTCAATTACCGCCTCTGCCACCGAGGAATTGGCCTTCTCTAGAAGACTCTTCCCCACAG
GAAGCCCCAAGTCAGCCGCCCGCTACATACTACTGGTAAATAAAGTTAAAAATAAGAAGACT
GTTTCCATCCCAAGCTACATTGAGCCTGAAGATGACTATGACGATGTTGAAATCCCTGCAAAT
ACTGAAAAAGCATCATTTTGAACAGCCATTTCTTCTTTTTTGGCAAAACTGAAGAGGGTTCAC
ACAACTTATTTTAAAACAATCAAGAATGGTTGAACTTCAGTAGGTCTCTGGGCCCTGAAAGCC
AGTGGTGATTTTATGAAGCTCTATAAGATAAAGCACTTCCCAAACCTTAGATGAAGACACCCC
TGCGATCGGATGACTGCAGCCAGAGGAGACACATGGGTGCTCGGCTCTGAGGACTTAGAGGGG
TCAGCCTTGTGCTGTTGAGGAACTTTCCATGGGAAGGACCACGGGGCTCCATGGCTCCCACC
TGTGGGAACTACTCATTTCTTGGCATTCTTCCCCCTTCATTCCCTTTGGTTTGCATGGTTC
TGAGTGATATTAAATCTCAGCATTTGGTTGTGCAAAAAAAAAA

[illegible]

FIGURE 580

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA144844
><subunit 1 of 1, 145 aa, 1 stop
><MW: 16618, pI: 5.26, NX(S/T): 1
MDTFTVQDSTAMSWWRNFWIILAVAIIVSVGLGLILYCVCKWQLRRGKKWEIAKPLKH
KQVDEEKMYENVLNE SPVQLPPLPPRNWPSLEDSSPQEAPSQPPATYSLVNKVKNKKTVS
IPSYIEPEDDYDDVEIPANTEKASF
```

Important features of the protein:

Signal peptide:

Amino acids 1-35

Tyrosine kinase phosphorylation site:

Amino acids 61-70

Amidation site:

Amino acids 48-52

FIGURE 581

GGCCGCCTCCGCGGGGCTGTGGGAAGCTTGGGCTGTCCCAGGACCGTCAGTCTCCTCCTCTGA
CCCTCCCTTTCCCTTGTGTGTAGGGCCGCGTCCCACCCCCACCTCGCCGGAGTCCGGGGCG
GCCCCGGTGTCCCTCCGAGCCTGCTGCACTCCACGTCCCCCTACCAGGGCTCCAGCCCCCAG
GGAAATCTCCGACCAGGCCCCGCCAGGAGCCAGATCCAGGCTCCTGGAAGAACCATGTCCGGC
AGCTACTGGTCATGCCAGGCACACACTGCTGCCCAAGAGGAGCTGCTGTTTGAATTATCTGTG
AATGTTGGGAAGAGGAATGCCAGAGCTGCCGGCTGAAAATTACCCAACCAAGAGAAATCTGCAGG
ATGGACTTTCTGGTCTCTTCTTGTTCACCTGGCTTCGGTGCTGATGGGTCTTGTTCCTTATC
TGCGTCTGCTCGAAAACCCATAGCTTGAAAGGCCTGGCCAGGGGAGGAGCACAGATATTTTCC
TGTATAATTCCAGAATGTCTTCAGAGAGCCGTGCATGGATTGCTTCATTACCTTTTCCATACG
AGAAACCACACCTTCATTGTCCTGCACCTGGTCTTGCAAGGGATGGTTTATACTGAGTACACC
TGGGAAGTATTTGGCTACTGTCAGGAGCTGGAGTTGTCCTTGCAATTACCTTCTTCTGCCCTAT
CTGCTGCTAGGTGTAAACCTGTTTTTTTTTACCCTGACTTGTGGAACCAATCCTGGCATTATA
ACAAAAGCAAATGAATTATTATTTCTTCATGTTTATGAATTTGATGAAGTGATGTTTCCAAAG
AACGTGAGGTGCTCTACTTGTGATTTAAGGAAACCAGCTCGATCCAAGCACTGCAGTGTGTGT
AACTGGTGTGTGCACCGTTTCGACCATCACTGTGTTTGGGTGAACAACCTGCATCGGGGCCTGG
AACATCAGGTACTTCCTCATCTACGTCTTGACCTTGACGGCCTCGGCTGCCACCGTCGCCATT
GTGAGCACCACCTTTCTGGTCCACTTGGTGGTGATGTCAGATTTATACCAGGAGACTTACATC
GATGACCTTGGACACCTCCATGTTATGGACACGGTCTTTCTTATTACGTACCTGTTCTGACT
TTTCCACGGATTGTCTTCATGCTGGGCTTTGTCGTGGTTCTGAGCTTCCTCCTGGGTGGCTAC
CTGTTGTTTGTCTGTATCTGGCGGCCACCAACCAGACTACTAACGAGTGGTACAGAGGTGAC
TGGGCCTGGTGCCAGCGTTGTCCCCTTGTGGCCTGGCCTCCGTCAGCAGAGCCCCAAGTCCAC
CGGAACATTCACTCCCATGGGCTTCGGAGCAACCTTCAAGAGATCTTTCTACCTGCCTTTCCA
TGTCATGAGAGGAAGAAACAAGAA**TGA**CAAGTGTATGACTGCCTTTG

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 582

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA144857
><subunit 1 of 1, 344 aa, 1 stop
><MW: 39787, pI: 7.44, NX(S/T): 2
MDFLVLFIFYLASVLMGLVLICVCSKTHSLKGLARGGAQIFSCIPECLQRAVHGLLHYL
FHTRNHTFIVLHLVLQGMVYTEYTWEVFGYCYQELLSLHYLLLPYLLLGVLNFFFTLTG
TNPGIITKANELLFLHVEYEFDEVMFKNVRCSTCDLRKPARSKHCSVCNWCVHRFDHHCV
WVNNCIGAWNIRYFLIYVLTLTASAATVAIVSTTFLVHLVVMSDLYQETYIDDLGHLHVM
DTVFLIQYLFLLTFPRIVFMLGFVVVLSFLLGGYLLFVLYLAATNQTTNEWYRGDWAWCQR
CPLVAWPPSAEPQVHRNIHSHGLRSNLQEIFLPAFPCHERKKQE
```

Important features of the protein:

Signal peptide:

Amino acids 1-29

Transmembrane domains:

Amino acids 100-116;201-217;256-275

N-glycosylation sites:

Amino acids 65-69; 284-290

N-myristoylation sites:

Amino acids 32-38; 77-83; 120-126; 322-328

Cell attachment sequence:

Amino acids 292-298

DHHC zinc finger domain:

Amino acids 140-204

FIGURE 583

CCGCGGAAGCTGGCAGGCGTTTCAGAGCGTCAGAGGCTGCGGATGAGCAGACTTGGAGGACTCCAGGCCAGAGACT
 AGGCTGGGCGAAGAGTCGAGCGTGAAGGGGGCTCCGGGCCAGGGTGACAGGAGGCGTGTCTTGAGAGGAAGAAGTT
 GACGGGAAGGCCAGTGCGACGGCAAATCTCGTGAACCTTGGGGGACGA**ATG**CTCAGGATGCGGGTCCCCGCCCTC
 CTCGTCTCTCTCTCTGCTTCAGAGGGAGAGCAGGCCCGTCGCCCATTTCTGCAACAGCCAGAGGACCTGGTG
 GTGCTGTGGGGGAGGAAGCCCGGCTGCCGTGTGCTCTGGGCGCCTACTGGGGGCTAGTTCACTGGACTAAGAGTGGG
 CTGGCCCTAGGGGGCCAAAGGGACCTACCAGGGTGGTCCCGTACTGGATATCAGGGAATGCAGCCAATGGCCAG
 CATGACCTCCACATTAGGCCCGTGGAGCTAGAGGATGAAGCATCATATGAATGTCAGGCTACACAAGCAGGCCCTC
 CGCTCCAGACCAGCCCAACTGCACGTGCTGGTCCCCCAGAAGCCCCCAGGTGCTGGGCGGCCCCCTCTGTGTCT
 CTGGTTGCTGGAGTTCTCTGCGAACCTGACATGTGCGAGCCGTGGGGATGCCCGCCCTACCCCTGAATTGCTGTGG
 TTCCGAGATGGGGTCTCTGTTGGATGGAGCCACCTTTTCATCAGACCCTGCTGAAGGAAGGGACCCCTGGGTGAGTG
 GAGAGCACCTTAACCTGACCCCTTTTCAGCCATGATGATGGAGCCACCTTTGTCTGCCGGGCCCCGAGCCAGGCC
 CTGCCCACAGGAAGAGACACAGCTATCACACTGAGCCTGCAGTACCCCCCAGAGGTGACTCTGTCTGCTTCGCCA
 CACACTGTGCAGGAGGGAGAGAAGGTCATTTTCTGTGCCAGGCCACAGCCAGCCTCCTGTACAGGCTACAGG
 TGGGCAAAGGGGGCTCTCCGGTGTCTCGGGGCCCGCGGGCCAAAGTTAGAGGTCGTGGCAGACGCCCTCGTTCCCTG
 ACTGAGCCCGTGTCTCTGCGAGGTGAGCAACGCGCTGGGTAGCGCCAACCGCAGTACTGCGCTGGATGTGCTGTTT
 GGGCCGATTCTGCAGGCAAAGCCGGAGCCCGTGTCCGTGAGCGTGGGGGAAGACGCTTCTTTCAGCTGCGCCTGG
 CGCGGGAACCCGCTTCCACGGGTAACCTGGACCCCGCGCGGTGGCGCGCAGGTGCTGGGCTCTGGAGCCACACTG
 CGTCTTCCGTGCGTGGGGCCCCGAGGACGCAGGCGACTATGTGTGCAGAGCTGAGGCTGGGCTATCGGGCCTGCGG
 GGCGGCGCCGCGGAGGCTCGGCTGACTGTGAACGCTCCCCAGTAGTGACCGCCCTGCACTCTGCGCCTGCCCTTC
 CTGAGGGGCCCCGCTCGCTCCAGTGTCTGGTTTTCTCGCTCTCCCCCCCCAGATGCCGTGGTCTGGTCTTGGGAT
 GAGGGCTTCTTGGAGGCGGGGTGCGAGGGCCGTTCTGTTGGAGACATTCCCTGCCCCAGAGAGCCGCGGGGA
 CTGGGTCCGGGCTGATCTCTGTGCTACACATTTCTGGGGACCCAGGAGTCTGACTTTAGCAGGAGCTTTAACTGC
 AGTGCCCGGAACCGGCTGGGCGAGGGAGGTGCCAGGCCAGCCTGGGCCGTAGAGACTTGCTGCCCACTGTGCGG
 ATAGTGGCCGGAGTGGCCCGTGCCACCACAACCTCTCCTTATGGTCATCACTGGGGTGGCCCTCTGCTGCTGGCGC
 CACAGCAAGGCCCTCAGCCTCTTTCTCCGAGCAAAGAACCTGATGCGAATCCCTGGCAGCAGCGACGGCTCCAGT
 TCACGAGGTCTGAAGAAGAGGAGACAGGCAGCCGCGAGGACCGGGGCCCATTTGTGCACACTGACCACAGTGAT
 CTGGTTCTGGAGGAGGAAGGGACTCTGGAGACCAAGGACCCCAACCAACGGTTACTACAAGGTCCGAGGAGTCAGT
 GTGAGCCTGAGCCTTGGCGAAGCCCTGGAGGAGGTCTCTTCTGCCACCACCCTCCCCCTTGGGCCCCCAGGG
 ACCCTTACCTTCTATGACTTCAACCCACACCTGGGCATGGTCCCCCCTGCAGACTTTACAGAGCCAGGGCAGGC
 TATCTCACCACACCCCAACCTCGAGCTTTTACCAGCTACATCAAACCCACATCCTTTGGGCCCCCAGATCTGGCC
 CCGGGACTCCCCCTTCCCATATGCTGCCTTCCCCACACCTAGCCACCCCGCTCTCCAGACTCACGT**GTGA**CAT
 CTTTCCAATGGAAGAGTCTGGGATCTCCAACCTTGCCATAATGGATTGTCTGATTTCTGAGGAGCCAGGACAAG
 TTGGCGACCTTACTCCTCCTCAAACTGAACACAAGGGGAGGGAAAGATCATTACATTTGTGAGGAGCATTTGTATA
 CAGTCAGCTCAGCCAAAGGAGATGCCCAAGTGGGAGCAACATGGCCACCCAATATGCCACCTATTCCCCGGTG
 TAAAAGAGATTCAAGATGGCAGGTAGGCCCTTTGAGGAGAGATGGGGACAGGGCAGTGGGTGTTGGGAGTTTGGG
 GCCGGGATGGAAGTTGTTTCTAGCCACTGAAAGAAGATATTTCAAGATGACCATCTGCATTGAGAGGAAAGGTAG
 CATAGGATAGATGAAGATGAAGAGCATACCAGGCCCAACCTGGCTCTCCCTGAGGGGAACCTTTGCTCGGCCAAT
 GGAAATGCAGCCAAGATGGCCATATACTCCCTAGGAACCAAAATGGCCACCATCTTGATTTTACTTTCTTAAA
 GACTCAGAAAGACTTGGACCAAGGAGTGGGGATACAGTGAGAATTACCACTGTTGGGGCAAATATTGGGATAA
 AAATATTTATGTTTAATAATAAAAAAAGTCAAAGAGAAAAAAA

589/615

FIGURE 584

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA145841
><subunit 1 of 1, 708 aa, 1 stop
><MW: 75093, pI: 6.65, NX(S/T): 3
MLRMRVPALLVLLFCFRGRAGPSPHFLQQPEDLVLLGEEARLPCALGAYWGLVQWTKSG
LALGGQRDLPGWSRYWISGNAANGQHDLHIRPVELEDEASYECQATQAGLRSRPAQLHVL
VPPEAPQVLGGPSVSLVAGVPANLTCRSRGDARPTPELLWFRDGVLLDGATFHQTLLKEG
TPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTAITLSLQYPPEVTLSASPHTVQE
GEKVIFLCQATAQPPVTGYRWAKGGSPVLGARGPRLEVADASFLETPVSCEVSNVAVGSA
NRSTALDVLFGPILQAKPEPVSDVGEDASFSCAWRGNPLPRVTWTRRGGAQVLGSGATL
RLPSVGPEDAGDYVCRAEAGLSGLRGGAAEARLTVNAPPVVTALHSAPAFLRGPARLQCL
VFASPAPDAVVSWEDEGFLEAGSQGRFLVETFPAPESRGGLGPGLISVLHISGTQESDFS
RSFNCSARNRLGEGGAQASLGRRDLLPTVRIVAGVAAATTTLLMVITGVALCCWRHASKAS
ASFSEQKNLMRIPGSSDGSSSRGPPEEEETGSREDRGPIVHTDHSDLVLEEEGTLETKDPT
NGYYKVRGVSVSLSLGEAPGGGLFLPPPSPLGPPGTPTFYDFNPHLGMVPPCRLYRARAG
YLTPHPRAFTSYIKPTSFGPPDLAPGTPPFYAAFPTPSHPRLQTHV

Important features of the protein:

Signal peptide:

Amino acids 1-20

Transmembrane domain:

Amino acids 511-531

N-glycosylation sites:

Amino acids 143-147;301-305;484-488

N-myristoylation sites:

Amino acids 48-54;60-66;79-85;139-145;180-186;183-189;355-361;383-389;
387-393;460-466;473-479;494-500;495-501;514-520;528-534;
554-560;592-598;608-614

Amidation site:

Amino acids 500-504

Cell attachment sequence:

Amino acids 149-152

Multicopper oxidases signature 1:

Amino acids 445-466

Immunoglobulin domain:

Amino acids 326-377

20251019 10:55:00

FIGURE 585

GCCCCGCTGAGGAAGCCGTGTGCCTGGGATGCCAAGAGCCAGAGAATGGATCTTCTCCGAGTG
GGGACATTGCTGACAATCCCGGCTTCCCGAGGCGGCTAAGAACAGGCAGTTTGTGTCCGCTGG
CTGCAGATACCCAGAGGCACAAAGAGACCGAAGCCACCCGGAGGGACCCACGGACGGACAGAT
GGTAGGCGCGAACCCGAGAGGACCGGCGGAGGCTGAGCACCGAGAGCCGCCAAGGAAGAGAAA
CTAACCACAGCCAAGTTACCCCGCCGGCTTTCTTCTCGCTGCGCTAAGGAATGAAACCCTTCCA
GCTCGATCTGCTCTTCTGCTGCTTCTTCTCTTCTCAGTCAAGAGCTGGGCCTCCAGAAGAGAGG
ATGCTGTCTGGTGCTGGGCTACATGGCCAAGGACAAGTTTTCGGAGAATGAATGAAGGCCAAGT
CTATTCCTTCAGCCAGCAGCCCCAGGACCAGGTGGTGGTGTCTGGGACAGCCAGTGACGCTACT
TTGCGCCATCCCCGAATACGATGGCTTCGTTCTGTGGATCAAGGACGGCTTGCGTCTGGGTGT
GGGCAGGGACCTCTCAAGTTACCCACAGTACCTGGTGGTAGGGAACACCTGTCAGGGGAGCA
CCACCTGAAGATCCTGAGGGCAGAGCTGCAAGACGATGCGGTGTACGAGTGCCAGGCCATCCA
GGCCGCCATCCGCTCCCGCCCCGCACGCCTCACAGTCTGGTGCCGCCTGATGACCCCGTCAT
CCTGGGGGGCCCTGTGATCAGCCTGCGTGCAGGGGGACCTCTCAACCTCACCTGCCACGCAGA
CAATGCCAAGCCTGCAGCCTCCATCATCTGGTTGCGAAAGGGAGAGGTCATCAATGGGGCCAC
CTACTCCAAGACCCTGCTTCGGGACGGCAAGCGGGAGAGCATCGTCAGCACCCCTCTTCATCTC
CCCTGGTGACGTGGAGAATGGCCAGAGCATCGTGTGTCGTGCCACCAACAAAGCCATCCCCGG
AGGAAAGGAGACGTCGGTCACCATTGACATCCAGCACCCCTCCACTGGTCAACCTCTCGGTGGA
GCCACAGCCAGTGCTGGAGGACAACGTCGTCACTTTCCACTGCTCTGCAAAGGCCAACCAGC
TGTCACCCAGTACAGGTGGGCCAAGCGGGGCCAGATCATCAAGGAGGCATCTGGAGAGGTGTA
CAGGACCACAGTGGACTACACGTACTTCTCAGAGCCCGTCTCCTGTGAGGTGACCAACGCCCTG
GGCAGCACCAACCTCAGCCGCACGGTTGACGTCTACTTTGGGCCCCGGATGACCACAGAACCC
CAATCCTTGCTCGTGGATCTGGGCTCTGATGCCATCTTCAGCTGCGCCTGGACCGGCAACCCA
TCCCTGACCATCGTCTGGATGAAGCGGGGCTCCGGAGTGCTCCTGAGCAATGAGAAGACCCTG
ACCCTCAAATCCGTGCGCCAGGAGGACGCGGGCAAGTACGTGTGCCGGGCTGTGGTGCCCCGT
GTGGGAGCCGGGGAGAGAGAGGTGACCCTGACCGTCAATGGACCCCCCATCATCTCCAGCACC
CAGACCCAGCACGCCCTCCACGGCGAGAAGGGCCAGATCAAGTGCTTCATCCGGAGCACGCCG
CCGCCGGACCGCATCGCCTGGTCTGGAAGGAGAACGTTCTGGAGTCGGGCACATCGGGGCGC
TATACGGTGAGACCATCAGCACCGAGGAGGGCGTCATCTCCACCCTGACCATCAGCAACATC
GTGCGGGCCGACTTCCAGACCATCTACAACCTGCACGGCCTGGAACAGCTTTCGGCTCCGACACT
GAGATCATCCGGCTCAAGGAGCAAGGTTTCGGAAATGAAGTCGGGAGCCGGGCTGGAAGCAGAG
TCTGTGCCGATGGCCGTCTCATTTGGGGTGGCCGTAGGAGCTGGTGTGGCCTTCCTCGTCCTT
ATGGCAACCATCGTGGCGTTCTGCTGTGCCCCGTTCCAGAGAAGTACGGGAGGGAGATCCGGG
ATCTCAGGGAGGGGGACAGAGAAAAAGGCCAGGCTTAGGCTGCCCCGAGAGCAAGTAAGCAG
GAGTGCAATGAACAGGGGTCCTAACAGTGCTGTGAGCTCCTGGGGCAGGGAGTGGGTCTGATG
CATCGGTGTATGTGAGCCTGGGCAACATGGCGCCTGGCAGAGTGGGCGCTAGGCTGAGGTTGA
CCTGGACTAGACTGAACTTCATCTGCAGGGCAGCCAGCATTTTGGATTGAACACATAGCTCTT
TCAGTCAGGAACTGTACAGAAAGATAGGGGGAAAAGCGGTTTGTGGTTTGATCCTTGCTCTAC
AAGAGCTGTTAGTCTAGAGAGACCCCATCTCTACAACAAAATAAAAATAAAGAGCTGCTAGTC
TCACCAGAAAAGCAGGTCACCTCACACAGCTGTGGGGGAGTGGGTGGGAAGCAATAAAGGAAT
TGCTTTGAGAAAACCTTAA

591/615

FIGURE 586

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA148004
><subunit 1 of 1, 600 aa, 1 stop
><MW: 65308, pI: 8.35, NX(S/T): 5
MKPFQLDLLFVCFLLFSQELGLQKRGCCCLVLGYMAKDKFRRMNEGQVYSFSQQPQDQVVV
SGQPVTLLCAIPEYDGFVLWIKDGLALGVGRDLSSYPQYLVVGNHLSGEHHLKILRAELQ
DDAVYECQAIQAAIRSRPARLTVLVPPDDPVILGGPVISLRAGDPLNLTCHADNAKPAAS
IIWLRKGEVINGATYSKTLRLDGKRESIVSTLFI SPGDVENGQSIVCRATNKAIPGGKET
SVTIDIQHPPLVNL SVEPQPVLEDNVVTFHCSAKANPAVTQYRWAKRGQIIKEASGEVYR
TTVDYTYFSEPV SCEVTNALGSTNLSRTVDVYFGPRMTTEPQSLLVDLGSDAIFSCAWTG
NPSLTIVWMKRGSGVVL SNEKTLTLKSVRQEDAGKYVCRAVVPRVGAGEREVTTLTVNGPP
IISSTQTQHALHGEKGQIKCFIRSTPPPDRIAWSKENVLES GTSGRYTVETISTEEGVI
STLTISNIVRADFQTIYNCTAWNSFGSDTEIIIRLKEQGSEM KSGAGLEAESVPMAVIIGV
AVGAGVAFLVLMATIVAFCCARSQRSTGGRSGISGRGTEKKARLRLPRRASKQECNEQGS

Important features of the protein:

Signal peptide:

Amino acids 1-17

Transmembrane domain:

Amino acids 534-555

N-glycosylation sites:

Amino acids 167-171;253-257;324-328;498-502

Glycosaminoglycan attachment sites:

Amino acids 523-527;574-578

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 204-208;370-374;588-592

Tyrosine kinase phosphorylation sites:

Amino acids 40-49;300-308;389-397

N-myristoylation sites:

Amino acids 45-51;62-68;84-90;103-109;192-198;236-242;
374-380;436-442;478-484;539-545;543-549;
568-574

Amidation site:

Amino acids 202-206

Leucine zipper pattern:

Amino acids 8-30

Downloaded from www.nrjonline.org

FIGURE 587

CAAAAAAGGAGCATGTCTTCATCCATGAGAGGCCTCGAAAAATCAATGGTATTTGCATTTCTC
CCAAGAAGGTTGCTTGCCAAAACCTTTCGGCCATTTTCTGCTTTCAGAGTGAGACAAAGTTCA
AAATGACAGTCTGTCAGCTCATTGAAGGCACTAGATACCCTGCCTGCAGGTACCACTATTCCC
CCACAGAGGGGTTTGTTCCTTGTCACTTGTGATGACTTGAGGCCAGATAGTTTCCTTGGCTATG
TTAAATAACTCAAGATCAGCTACCGAGTCTGAGATCTCTTCTCTCATGGCATTGGAGCTGGCT
GTGCCTGAGGCAGACCTGGACCGTGGACATGGGGCAATGCCTTGAGCGGAAGGGGAAGCCACT
GAATTTTGGGTGTCACCAGGTAAACAGAGCCCTCAGCATCTGAATAGAACTGAACAGGAACA
GAAGAGATTACACTACATCTGAGATGGAGACCTTTTCTCTGCTGCTGCTCAGCCTGGGCCTGG
TTCTTGCAGAAGCATCAGAAAAGCACAATGAAGATAATTAAAGAAGAATTTACAGACGAAGAGA
TGCAATATGACATGGCAAAAAGTGGCCAAGAAAAACAGACCATTGAGATATTAATGAACCCGA
TCCTGTTAGTTAAAAATACCAGCCTCAGCATGTCCAAGGATGATATGTCTTCCACATTACTGA
CATTCAGAAGTTTACATTATAATGACCCCAAGGGAAACAGTTCGGGTAAATGACAAAGAGTGTT
GCAATGACATGACAGTCTGGAGAAAAGTTTCAGAAGCAAACGGATCGTGCAAGTGGAGCAATA
ACTTCATCCGCAGCTCCACAGAAGTGATGCGCAGGGTCCACAGGGCCCCCAGCTGCAAGTTTGTA
CAGAATCCTGGCATAAGCTGCTGTGAGAGCCTAGAACTGGAAAATACAGTGTGCCAGTTCACT
ACAGGCAAACAATTCCCCAGGTGCCAATACCATAGTGTTACCTCATTAGAGAAGATATTGACA
GTGCTGACAGGTCATTCTCTGATGAGCTGGTTAGTTTGTGGCTCTAAGTTGTAAATCCCACAG
AGCTTTAGGACTAGGGTCTTACTAAAGAAGGACCTCTTCTTGTTTCAATTCTTGTTTAAACCTTT
CCTTAATATCTACTCTTTAGCACTATAGTGAACCTCTGATTATTTATTCTAACTGGAGGAGTG
AAAAATCCAAAATTGTGGATAATTCAATTAAAGTTATGACTGATACCG

593/615

FIGURE 588

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA149893
><subunit 1 of 1, 199 aa, 1 stop
><MW: 22427, pI: 6.46, NX(S/T): 3
METFPLLLLLSLGLVLAEASESTMKIIKEEFTDEEMQYDMAKSGQEKQTIEILMNPILLVK
NTSLSMSKDDMSSTLLTFRSLHYNDPKGNSGNDKECCNDMTVWRKVSEANGSCKWSNNF
IRSSTEVMRRVHRAPSCKFVQNPGISCCESLELENTVCQFTTGKQFPRCQYHSVTSLEKI
LTVLTGHSLMSWLVCGSKL

Important features of the protein:

Signal peptide:

Amino acids 1-16

N-glycosylation sites:

Amino acids 61-65;89-93;111-115

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 105-109

N-myristoylation sites:

Amino acids 12-18;88-94;144-150

Microbodies C-terminal targeting signal:

Amino acids 197-201

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 589

CAGTCCTGCCGGGACGGTGAGCGCATTCAGCACCTGGACAGCACCGCGGTTGCGCTGCCTCC
AGGGCGGCCCCGGGCTGCTCCTGCTCCGCAGAGCTACGCCCTCCCCCGGGTGCCCCGGACCC
TGCACTTGCCGCCGCTTTCTCTCGCGCTGCTCTGGACCTTGCTAGCCGGCTCTGCACCTCCCAG
AAGCCGTGGGCGCGCCGCTCAGCTGCTCCATCGCCTCACTTTCCAGGCTCGCGCCCGAAGCA
GAGCCATGAGAACCCCAGGGTGCTGGCGAGCCGCTAGCGCCATGGGCCCCGGCGAGGCGCTG
CTGGCGGGTCTCCTGGTGATGGTACTGGCCGTGGCGCTGCTATCCAACGCACTGGTGCTGCTT
TGTTGCGCTACAGCGCTGAGCTCCGCACTCGAGCCTCAGGCGTCTCCTGGTGAATCTGTCT
CTGGGCCACCTGCTGCTGGCGGCGCTGGACATGCCCTTCACGCTGCTCGGTGTGATGCGCGGG
CGGACACCGTCGGCGCCCCGGCGCATGCCAAGTCATTGGCTTCCTGGACACCTTCCTGGCGTCC
AACGCGGCGCTGAGCGTGCGGCGCTGAGCGCAGACCAGTGGCTGGCAGTGGGCTTCCCACTG
CGCTACGCCGGACGCCTGCGACCGCGCTATGCCGGCCTGCTGCTGGGCTGTGCCTGGGGACAG
TCGCTGGCCTTCTCAGGCGCTGCACTTGGCTGCTCGTGGCTTGGCTACAGCAGCGCCTTCGCG
TCCTGTTTCGCTGCGCCTGCCGCCCGAGCCTGAGCGTCCGCGCTTCGCAGCCTTCACCGCCACG
CTCCATGCCGTGGGCTTCGTGCTGCCGCTGGCGGTGCTCTGCCTCACCTCGCTCCAGGTGCAC
CGGGTGGCACGCAGACACTGCCAGCGCATGGACACCGTCACCATGAAGGCGCTCGCGCTGCTC
GCCGACCTGCACCCAGTGTCGGCGAGCGCTGCCTCATCCAGCAGAAGCGGCGCCGCCACCGC
GCCACCAGGAAGATTGGCATTGCTATTGCGACCTTCCTCATCTGCTTTGCCCCGTATGTCATG
ACCAGGCTGGCGGAGCTCGTGCCCTTCGTACCGTGAACGCCAGTGGGGCATCCTCAGCAAG
TGCCTGACCTACAGCAAGGCGGTGGCCGACCCGTTACGTA CTCTCTGCTCCGCCGGCCGTTT
CGCCAAGTCCTGGCCGGCATGGTGCACCGGCTGCTGAAGAGAACCCCGCGCCAGCATCCACC
CATGACAGCTCTCTGGATGTGGCCGGCATGGTGCACCAGCTGCTGAAGAGAACCCCGCGCCCA
GCGTCCACCCACAACGGCTCTGTGGACACAGAGAATGATTCTGCCTGCAGCAGACACACTTGA
GGGCCTGGCAGGGCTCATCGCCCCACCTTCTAAGAAGCCCTGTGGAAAGGGCACTGGCCCTG
CCACAGAGATGCCACTGGGGACCCCCAGACACCAGTGGCTTGACTTTGAGCTAAGGCTGAG

595/615

FIGURE 590

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA149930
><subunit 1 of 1, 363 aa, 1 stop
><MW: 39332, pI: 10.42, NX(S/T): 3
MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSaelRTRASGVLLVNLSLGHLLLAALDM
PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP
RYAGLLLGCawGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG
FVLPLAVLCLTSLQVHRVARRHQCQRMdVTMkALALLADLHPSVRQRCLIQKRRRHRA
RKIGIAIATFLICFAPYVMTRLAEVFPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRP
FRQVLagmvHRLlKRTPrPASTHDSSLDVagmvHQLlKRTPrPASTHNGSVDTENDSCLQ
QTH

Important features of the protein:

Signal peptide:

Amino acids 1-24

Transmembrane domains:

Amino acids 46-60;85-103;130-150;175-192;246-264

N-glycosylation sites:

Amino acids 47-51;348-352;355-359

Tyrosine kinase phosphorylation site:

Amino acids 286-295

N-myristoylation sites:

Amino acids 66-72;124-130;128-134;132-138;139-145;244-250;
349-355

G-protein coupled receptor proteins:

Amino acids 72-112

7 transmembrane receptor (rhodopsin family):

Amino acids 22-294

Downloaded from www.nrjonline.org

AAC**ATG**AGCTGCGGCGCCTGGGCTGCTCGTCTGGCTGCTCGTGCTCCGGCTGCCCTGGCGGGTG
CCGGGCCAGCTGGACCCCAGCACTGGCCGGCGGTTCTCGGAGCACAACTCTGCGCGGACGAC
GAATGCAGCATGATGTACCGCGGTGAGGCTCTTGAAGATTTACAGGCCCGGATTGTCTGTTTT
GTGAATTTTTAAAAAAGGTGATCCTGTATATGTTTACTATAAACTGGCAAGAGGATGGCCTGAA
GTTTGGGCTGGAAGTGTGGACGCACTTTTGGATATTTTCCAAAAGATTTAATCCAGGTAGTT
CATGAATATACCAAAGAAGAGCTACAAGTTCCAACAGATGAGACGGATTTTGTGTTTGTGTTTGTAT
GGAGGAAGAGATGATTTTCATAATTATAATGTAGAAGAACTTTTAGGGTTTTTGGAACTGTAC
AATTCTGCAGCTACAGATTCTGAGAAAAGCTGTAGAAAAAACTTTACAGGATATGGAAAAAAAC
CCTGAATTATCTAAGGAAAGGGAACCTGAACCTGAACCAGTAGAAGCCAACTCAGAGGAAAGT
GATAGTGTATTCTCAGAAAAACTGAGGATCTTCAGGAACAGTTTACAACCTCAGAAGCACCAC
TCCCATGCAAAACAGCCAAGCAAATCATGCTCAGGGAGAGCAGGCTTCATTTGAATCTTTTGAA
GAAATGCTGCAAGATAAACTAAAAGTGCCAGAAAAGTGAAAACAACAAAACCAGCAATAGTTCT
CAGGTCTCAAATGAACAGGATAAGATTGATGCCTATAAACTTTTGAAAAAAGAAATGACTCTA
GACTTGAAAACCAAATTTGGCTCAACAGCTGATGCACTTGTATCTGATGATGAGACAACCAGA
CTCGTTACTTCAATTAGAAGATGATTTTGTATGAGGAATTGGATACTGAGTATTATGCAGTTGGA
AAGGAAGATGAGGAGAACCAAGAAGACTTTGATGAGTTGCCATTACTTACCTTTACAGATGGG
GAAGATATGAAAACCTCCAGCAAAGTCTGGCGTTGAGAAATATCCAACAGATAAAGAGCAGAAT
TCAAATGAAGAGGACAAGGTTTCACTAACTGTGCCCCCTGGCATCAAAAATGATGATAAAAAT
ATACTAACAACTGGGGGGACACTATCTTCTCTATTGTACAGGAGGTGAAGAAACAAGAGAT
ACGATGGATTTAGAGAGCTCTAGTTTCAAGGGAAGAAAAAGAAGATGATGATGATGCATTAGTC
CCAGATAGCAAAACAGGGGAAACCAAGTCAAGTCAAGCAACAGATTATAGTGACCCTGACAATGTAGAT
GATGGTCTTTTTTATTGTAGACATTCTTAAACAAATAATGACAAAGAAGTAAACGCAGAACAT
CACATTAAAGGAAAAGGGAGGGGAGTTCAAGGAATCCAAGAGGGGCCTGGTACAAGATGAGACA
GAATTAGAGGATGAAAATCAAGAAGGCTTTTAAACAGAGCCCATAAAACTAT**TGA**CCTCTGAGG
TTTCATTGGAAAGAAAAGTGTACTGTGCATTATCCATTACAGTAAAGGATTTTATTGGCTTCAA
AATCCAAAAGTTTATTTTAAAAAGTTTGTGTTGTAGAACTAAGCTGCCTTGGCAGTGTGCATTT
TTGAGCCAAAACAAATCAAAAATGTCAATTTCTTCCCTAAATAAAAATCACCTTTTTAAGCTAGAG
CGTCCTTACAACCTTTGAAATGTGCAATAAAGAATACCTGTGTTTTAGCTAATGTAGCATATGT
AATTGCAAAATGATTTAGAATGTCATGAAAAATATGAACATTTCTGTGGAAATGCTTTAAGA
ACATGTATTTCCATTATCCTATTTTTTAGTGTACACCAGCTGAATACGGAGCAATGGTGTGTTAT
AAGCGTTTTTTTTAACTATCTGGTCACAAAGACTGTTACGCTAAAAATGTTTACTAAAAGATC
ACTAAACTATCTCCCTCTTGCTGAAGTTCTTTGTAGTAATAGCTCATAAAAATTTGTTTATT
AATATTTAAAAA

597/615

FIGURE 592

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA150157
><subunit 1 of 1, 499 aa, 1 stop
><MW: 56471, pI: 4.31, NX(S/T): 2
MAAAPGLLVWLLVLRPWRVPGQLDPSTGRRFSEHKLCADDECSMMYRGEALEDFTGPDC
RFVNFKKGDVPYVYYKLARGWPEVWAGSVGRTFGYFPKDLIQVVHEYTKHEELQVPTDET
FVCFDGGRRDDFHNYNVEELLGFLELYNSAATDSEKAVEKTLQDMEKNPELSKEREPEPEP
VEANSEESDSVFSENTEDLQEQFTTQKHHSHANSQANHAQGEQASFESFEEMLQDKLKVP
ESENKTSNSSQVSNEQDKIDAYKLLKKEMTLDLKTKFGSTADALVSDDETTRLVTSLED
DFDEELDTEYYAVGKEDEENQEDFDELPLLTFTDGEDMKTPAKSGVEKYPTDKEQNSNEE
DKVQLTVPPGIKNDKINILTTWGDITFSIVTGGEETRTMDLESSSSEEEKEDDDDALVP
DSKQKGKQSATDYSDDPNVDDGLFIVDIPKTNNDKEVNAEHHIKGKGRGVQESKRGLVQD
ETELDENQEGFKTEPIKL

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-glycosylation sites:

Amino acids 245-249;249-253

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 30-34

Tyrosine kinase phosphorylation site:

Amino acids 66-72

N-myristoylation sites:

Amino acids 392-398;469-475

Amidation site:

Amino acids 28-32

Aminoacyl-transfer RNA synthetases class-II signature 1:

Amino acids 47-70

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

598/615

FIGURE 593

GGGCCAGTAGAGTGTGTCTGGGTCTGAGTGACTACATCAAAGCTCCCAGCCTTGAAAAAC
ACATGCTGTTCCCAGGCCTCAAGATATTGAAACATTAATTAGATAATTTAAAGTAGCGTTTTTC
TTCTACAATGCTCTGAAGAAGTGACCTACGCGACACTCACATTTTCAGGATTCTGCTGGAGCAAG
GAATAACCGAGATGGAAATAACCTAAGAAAAAGAGGGCATCCAGCTCCATCTCCCATTGCGG
TCATGCTGCTCTGGGTCTGGTAACTCTTTGCCTGATGTTGCTGATTGGGCTGGTGACGTTGGG
GATGATGTTTTTGCAGATATCTAATGACATTAACCTCAGATTCAGAGAAATTGAGTCAACTTCA
GAAAACCATCCAACAGCAGCAGGATAACTTATCCCAGCAACTGGGCAACTCCAACAACCTTGTC
CATGGAGGAGGAATTTCTCAAGTCACAGATCTCCAGTCTACTGAAGAGGCAGGAACAAATGGC
CATCAAACCTGTGCCAAGAGCTAATCATTACACTTCAGACCACAGATGTAATCCATGTCCTAA
GATGTGGCAATGGTACCAAAAATAGTTGCTACTATTTTACAACAAATGAGGAGAAAACCTGGGC
TAACAGTAGAAAGGACTGCATAGACAAGAACTCCACCCTAGTGAAGATAGACAGTTTGGAAGA
AAAGGATTTTCTTATGTCACAGCCATTACTCATGTTTTTCGTTCTTTTGGCTGGGATTATCATG
GGACTCCTCTGGCAGAAGTTGGTTCTGGGAAGATGGCTCTGTTCCCTCTCCATCCTTGTACGT
CTCTAACTATTGAGGGGTAAACACAAGCTTTCCATGGAATCCTGGGAAAATTAATAATGATTGT
GAGAATTATAAATACAGACATAAAAAGAGGAGTACAACATACTGAGAAAAGAGCTCCAGTAAC
AAATATTGAAAGGAGATTTAGTACTAAAGAACTTGACCAGATCAATGGATCCAAAGGATGTGC
TTATTTTCAAAAAGGAAATATTTATATTTCTCGCTGTAGTGCTGAAATTTTTTGGATTGCGA
GAAGACAGCTGCCCCAGTGAAGACTGAGGATTTGGATTAGTATGCTTCTTCCAAATTCTCCAA
GAAGTAAGAGACTTGTGAGTAAGCTCATATGAGGAAAGAGGAAACTACGGTACCAGAGCAAGG
GCGAATTCTGCA

2025-04-04 10:55:00

599/615

FIGURE 594

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA150163
><subunit 1 of 1, 232 aa, 1 stop
><MW: 26754, pI: 5.80, NX(S/T): 3
MSEEVITYATLTFQDSAGARNNRDGNLRLKRGHPAPSPIWRHAALGLVTLCLMLLIGLVTL
GMMFLQISNDINSDEKLSQLQKTIQQQODNLSQQLGNSNNLSMEEEFLLKSQISSLLKRQ
EQMAIKLCQELIIHTSDHRCNFCPKMWQWYQNSCYFYFTTNEEKTWANSRKDCIDKNSTLV
KIDSLEEKDFLMSQPLLMFSFFWLGLSWDSSGRSWFWEDGSVPSPSLYVSNY
```

Important features of the protein:

Transmembrane domain:

Amino acids 42-62

N-glycosylation sites:

Amino acids 91-95;101-105;176-180

N-myristoylation sites:

Amino acids 17-23;97-103

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 595

CGGACGCGTGGGGAAGAGGAGGAGGAGGAAGAAGACGTGGACAAGGACCCCCATCCTACCCAG
AACACCTGCCTGCGCTGCCGCCACTTCTCTTTAAGGGAGAGGAAAAGAGAGCCTAGGAGAACC
ATGGGGGGGCTGCGAAGTCCGGGAATTTCTTTTGCAATTTGGTTTCTTCTTGCCCTATGCTGACA
GCGTGGCCAGGCGACTGCAGTCACGTCTCCAACAACCAAGTTGTGTTGCTTGATACAACAAC
GTACTGGGAGAGCTAGGATGGAAAACATATCCATTAAATGGGTGGGATGCCATCACTGAAATG
GATGAACATAATAGGCCCATTACACATAACCAGGTATGTAATGTAATGGAACCAAACCAAAC
AACTGGCTTCGTACAAACTGGATCTCCCGTGATGCAGCTCAGAAAATTTATGTGGAAATGAAA
TTCACACTAAGGGATTGTAACAGCATCCCATGGGTCTTGGGGACTTGCAAAGAAACATTTAAT
CTGTTTTATATGGAATCAGATGAGTCCCACGGAATTAAATTCAAGCCAAACCAGTATACAAAG
ATCGACACAATTGCTGCTGATGAGAGTTTTACCCAGATGGATTGTTGGGTGATCGCATCCTCAA
CTCAACACTGAAATTCGTGAGGTGGGGCCTATAGAAAGGAAAGGATTTTATCTGGCTTTTCAA
GACATTGGGGCGTGCAATTGCCCTGGTTTCAGTCCGTGTTTTCTACAAGAAATGCCCTTCACT
GTTTCGTAACCTGGCCATGTTTCCTGATACCATTCCAAGGGTTGATTCTCCTCTTTGGTTGAA
GTACGGGGTTCTTGTGTGAAGAGTGCTGAAGAGCGTGACACTCCTAACTGTATTGTGGAGCT
GATGGAGATTGGCTGTTTCTTGGAAAGGTGCATCTGCAGTACAGGATATGAAGAAATTGAG
GGTTCTTGCCATGGAGCCTCAAAGGCCGCTGCTTCT**AG**TTGGCCATCTTGGCCCCACCCGA
AACAGTAACCTTTGAAGAATAAAAGAAAAAGCAAAGAGTAGCATTACTAAAATATTAAACGG
TTACATTTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

601/615

FIGURE 596

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA153579
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32368, pI: 5.32, NX(S/T): 0
MGGCEVREFLLQFGFFLPMLTAWPGDCSHVSNNQVVLLDTTTVLGELGWKTYPLNGWDAI
TEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWI SRDAAQKIYVEMKFTLRDCNSIPWVLGT
CKETFNLFYMESEDESHGIKFKPNQYTKIDTIAADESFTQMDLGDRILKLNTEIREVGPIE
RKG FYLAFQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCVKSA
EERDTPKLYCGADGDWLVPLGRCICSTGYEEIEGSCHGASKGRCF
```

Important features of the protein:

Signal peptide:

Amino acids 1-22

N-myristoylation sites:

Amino acids 192-198;274-280;278-284

Receptor tyrosine kinase class V signature 1:

Amino acids 192-209

Ephrin receptor ligand binding domain:

Amino acids 34-207

601/615

FIGURE 597

ACACTGGCCAAACAACAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAAGGACTCCCA
GGACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGG
ATTCCAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCCCTTGGGGGGGGG
CAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAG**ATG**CCTGTGCCCTGGTTCTT
GCTGTCTTTGGCACTGGGCGGAAGCCCAGTGGTCCTTTCTCTGGAGAGGCTTGTGGGGCCTCA
GGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTCTGGGACAGTGACATACTCTGCC
GCCTGGGGACATCGTGCCTGCTCCGGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGC
GGTGTCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC
CGTGCATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGG
GGTGGAGGAGCCTAGGAATGCCTCTCTCCAGGCCCCAAGTCGTGCTCTCCTTCCAGGCCTACCC
TACTGCCCCGCTGCGTCTTGCTGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTC
TGTGGGCTCTGTGGTATATGACTGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTC
CTATACTCAGCCCAGGTACGAGAAGGAACTCAACCACACACAGCAGCTGCCTGCCCTGCCCTG
GCTCAACGTGTCTCAGCAGATGGTGACAACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCA
CTTCGGCCTCTCCCTGTACTGGAATCAGGTCCAGGGCCCCCCCCAAAACCCCGGTGGCACAAAAA
CCTGACTGGACCGCAGATCATTAACCTTGAACCACACAGACCTGGTTCCCTGCCTCTGTATTCA
GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCTTCCAGGGAGGACCCCCG
CGCACACCAGAACCTCTGGCAAGCCGCCCGACTGCGACTGCTGACCTTGCAGAGCTGGCTGCT
GGACGCACCGTGCTCGCTGCCCCGAGAAGCGGCACTGTGCTGGCGGGCTCCGGGTGGGGACCC
CTGCCAGCCACTGGTCCCACCGCTTTCCTGGGAGAACGTCACTGTGGACAAGGTTCTCGAGTT
CCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTTCAAGGTGAACAGCTCGGAGAAGCTGCAGCT
GCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGACGATGTGCTACTGTTGGAGAC
ACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAAACCAGTGGCTGTACTTCACTACC
CAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTACAAGACCTGCAGTCAGG
CCAGTGCTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTGCCCCATGGACAAATA
CATCCACAAGCGCTGGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCGCTGCGCTTTCCCT
CATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAAACAGGACGTCCG
CTCGGGGGCGGGCCGCCAGGGGCCGCGCGGCTCTGCTCCTCTACTCAGCCGATGACTCGGGTTT
CGAGCGCCTGGTGGGCGCCCTGGCGTCGGCCCTGTGCCAGCTGCCGCTGCGCGTGGCCGCTAGA
CCTGTGGAGCCGTCTGAACTGAGCGCGCAGGGGCCCGTGGCTTGGTTTTCACGCGCAGCGGCG
CCAGACCCTGCAGGAGGGCGGCGTGGTGGTCTTGCTCTTCTCTCCCGGTGCGGTGGCGCTGTG
CAGCGAGTGGCTACAGGATGGGGTGTCCGGGGCCGGGGCGCACGGCCCGCACGACGCCTTCCG
CGCCTCGCTCAGCTGCGTGCTGCCCGACTTCTTGACAGGGCCGGGGCGCCCGGCAGCTACGTGGG
GGCCTGCTTCGACAGGCTGCTCCACCCGACGCCGTACCCGCCCTTTTCCGCACCGTGCCCGT
CTTCACACTGCCCTCCCAACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCG
TTCCGGGGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAG
CTACTTCCATCCCCCGGGGACTCCCGCGCCGGGACGCGGGGTGGGACCAGGGGCGGGGACCTGG
GGCGGGGGGACGGGACT**TAAATAAAGGCAGACGCTGTTTTTCTAAAAAA**

603/615

FIGURE 598

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA164625

><subunit 1 of 1, 705 aa, 1 stop

/><MW: 76970, pI: 6.00, NX(S/T): 9

MPVPWFLLSLALGRSPVVLSELRLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGP
VLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRN
ASLQAQVVLSFQAYPTARCVLLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYT
QPRYEKELNHTQQPLPALPWLNV SADGNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWH
KNLTGPQIIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLLT
LQSWLLDAPCSLPAEAALCWRA PGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCV
QVNSSEKLQLQECLWADSLGPLKDDVLLLETRGPQDNRS LCALEPSGCTSLPSKASTRA
ARLGEYLLQDLQSGQCLQLWDDDLGALWACPM D KYIHKRWALVWLACLLFAAALSLILL
LKKDHAKGWLRLLLKQDVRS GAAARGRAALLLYSADDSGFERLVGALASALCQLPLRVAV
DLWSRRELSAQGPVAFWFAQRRTLQEGGVVLLFSPGAVALCSEWLQDGVSGPGAHP
HDAFRASLS CVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFTLPSQLPDLG
ALQQPRAPRSGRLQERAEQVSRA LQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Important features of the protein:

Signal peptide:

Amino acids 1-20

Transmembrane domain:

Amino acids 453-473

N-glycosylation sites:

Amino acids 118-122;186-190;198-202;211-215;238-242;
248-252;334-338;357-360;391-395;

Glycosaminoglycan attachment site:

Amino acids 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site:

Amino acids 552-556

N-myristoylation sites:

Amino acids 107-113;152-158;319-325;438-444;516-522;612-618;
692-698;696-702;700-706

GGTCCCTTAATGGCAGCAGCCGCGCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGC
TGTCCGGCTGGTCCC CGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCA
TCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTT
TTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCACTCCCCTGGGGAAGAAACTAA
ATGTCACAACGGCCTGGAAAAGCACAGAACCAGTACTGAGAGAGGTGGTGGACATACTTACAG
AGCAACTGCGTGACATTCACTGTGGAGAATTACACACCCAAGGAACCCCTCACCTGCAGGCCAA
GGATGTCTTTGTGAGCAGAAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATG
GGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCA
GAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAA
TGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCTGGAGCCAA
GTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCACAGCCACCACCC
TCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCATCTGAAGGAGAGT
CCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCACGGTCTTGATCAAAC
TCGCCCTTCTGTCTGGCCAGCTGCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGAT
CATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCTTTTTGCCAACAAATTTTA
CCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCA
CTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTCTCTTTTTGTTTGGAAAA
TCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAAGTAAAATAATCACG
TTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTTTAAATTATTTAAT
AAGAAAAAATTTATATTAATGATTGTTTTCTTTAGTAATTTATTGTTCTGTACTGATATTTAA
ATAAAGAGTTCATTTCCCAAAAAAAAAAAAAAAAAAAAA

605/615

FIGURE 600

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFLH
YDCGNKTVTPVSP LGKKLNVT TAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQARMS
CEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGD
CIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

605/615
FIGURE 600

FIGURE 601

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCTT
GCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCTCA
CCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATGATG
GGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCCCGGC
GCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGACTTTGT
GCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACCAGCTCT
CCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCCCAAGAGT
TGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGCTGAAAAAC
TCTGTCTGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAGAACAATGGA
AATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAGGACTGTAAAT
ATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAACAAGAAGACCTGGAATTTG
CCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTTTTCGCCCTGACA
GTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACCTCTGAACTGTTCCATATTATAA
TAGATGTCACCAGCCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGGGATGATCTTCTCAA
AGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGA
GCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACTGATTTCGCCCTCTGCAACTACAAAT
AGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACATTGGGAAATGGAACATAA
TCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTGTTTCCTGTTTCAGGATCAC
CAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGC
CACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGC
TCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCATGTCTTCCTTACACTTGGTGGA
ATAAGAAACTTTTTTGAAGTAGAGGAAATACATTGAGGTAACATCCTTTTCTCTGACAGTCAAG
TAGTCCATCAGAAATTGGCAGTCACTTCCAGATTGTACCAGCAAATACACAAGGAATTCTTT
TTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCCCATCTGCCTTGT
CCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAGAATCTCAAATCTCAATGCCTTATAA
GCATTCTTCCTGTGTCCATTAAGACTCTGATAATTGTCTCCCCTCCATAGGAATTTCTCCCA
GGAAAGAAATATATCCCCATCTCCGTTTCATATCAGAACTACCGTCCCCGATATTCCCTTCAG
AGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCATCTGCACCTGTAATAGTTTCAGTTCCTA
TTTTCTTCCATTGACCCATATTTATACCTTTCAGGTACTGAAGATTTAATAATAATAAATGTA
AATACTGTGAAAAA

607/615

FIGURE 602

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCVLII
GLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRELY
NKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAASQSY
SEFFYSYWTGLLRPD SGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKDCKEL
KRCVCERRAGMVKPESLHVPPETLGEGD

607/615

608/615

FIGURE 603

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTCC
CAGCAAT**ATG**CATCTTGACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCCCT
GTCTGGATGGGCGGCCAGCGATGACCCATTGAGAAGGTCATTGAAGGGATCAACCGAGGGCT
GAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGAATCACGCATGC
CGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGCAAGGA
GTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGATCAACCA
TGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACGCTGCTGG
ACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCACCAGGCTGG
GAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGGAAAGGAAGT
GGAGAAGCTTGGCCAAGGTGCCCAACCATGCTGCTGGCCAGGCCGGAAGGAGCTGCAGAATGC
TCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGCAACCATCAAAG
CGGATCTTCCAGCCATCAAGGAGGGGGCCACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAA
CACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCATGCCCT**TAA**ACTGG
CATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTACATCAGCTGACATGACCTGGAGGGG
TTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTGGGATTGTGAATAAA
CTTGATACACCA

10092386.011503

609/615

FIGURE 604

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675
><subunit 1 of 1, 247 aa, 1 stop
><MW: 25335, pI: 7.00, NX(S/T): 0
MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAGR
EVEKVFNGLSNMGSHGTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHG VNNAAGQA
GKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQNAHN
GVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

00556 0100 98926001

610/615

FIGURE 605

GCGACGCGCGGCGGGGCGGCGGAGAGGAAACGCGGCGCCGGGCGGGCCCCGGCCCTGGAGATGG
TCCCCGGCGCCGCGGGCTGGTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCCACG
GCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTTCGATACATCT
TCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATTTCACC
TTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACTCAGCAACGGTTTCTTCATCCAGGACC
AGATTGCTCTGGTGGAGAGGGGGGCTGCTCCTTCTCTCCAAGACTCGGGTGGTCCAGGAGC
ACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTACGTGGAGA
TGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGGCCGAGACG
GCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTTCCATCCCAG
TCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTCTGGTAGAAGA
GTTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCCAGGAATTTTGC
TACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGAAAAGGGTTTGGG
CGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGGGCCCCCAAGGGTG
TCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTAGAACCCGAAACAAA
AGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACTCACCTGGCTCCAGCC
TCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGGAGTGGTTTAAAGAGCT
GGTGTTTGGGGACTCAATAAACCTCACTGACTTTTGTAGCAATAAAGCTTCTCATCAGGGTTG
CAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

20510-999999999

611/615

FIGURE 606

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532
><subunit 1 of 1, 188 aa, 1 stop
><MW: 21042, pI: 5.36, NX(S/T): 2
MVPGAAGWCCLVLWLPACVAAHGFRIHDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQI
HLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVII SDNAVDNDSFYV
EMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW
```

Signal peptide:
amino acids 1-20

606110-606500

FIGURE 607

GCATTTGCCACTGGTTGCAGATCAGGCGGACGAGGAGCCGGGAAGGCAGAGCC**ATGT**GGCTGC
 CCCCTGCTCTGCTCCTTCTCAGCCTCTCAGGCTGTTTCTCCATCCAAGGCCAGAGTCTGTGA
 GAGCCCCAGAGCAGGGGTCCCTGACGGTTCAATGCCACTATAAGCAAGGATGGGAGACCTACA
 TTAAGTGGTGGTGCCGAGGGGTGCGCTGGGATACATGCAAGATCCTCATTGAAACCAGAGGGT
 CGGAGCAAGGAGAGAAGAGTGACCGTGTGTCCATCAAGGACAATCAGAAAGACCGCACGTTCA
 CTGTGACCATGGAGGGGCTCAGGCGAGATGACGCAGATGTTTACTGGTGTGGGATTGAAAGAA
 GAGGACCTGACCTTGGGACTCAAGTGAAAGTGATCGTTGACCCAGAGGGAGCGGCTTCCACAA
 CAGCAAGCTCACCTACCAACAGCAATATGGCAGTGTTTCATCGGCTCCCACAAGAGGAACCACT
 ACATGCTCCTGGTATTTGTGAAGGTGCCCATCTTGCTCATCTTGGTCACTGCCATCCTCTGGT
 TGAAGGGGTCTCAGAGGGTCCCTGAGGAGCCAGGGGAACAGCCTATCTACATGAACTTCTCCG
 AACCTCTGACTAAAGACATGGCCACT**TAG**AGAGATGGATCTGCAGAGCCTTCTGCCCCTGGCC
 ACGTTTCCAGAAGAGACTCGGGCTGTGGAAGGAACATCTACGAGTCCTCGGGATGCAGTGACT
 GAGATAGGGGCCCTGGGCCTCCGCCCTGGCCTTGGAGCTGGTGGGCACCTCCCTGTTCTGCAC
 AGCTCAGGGACTTAGCCAGGTCTCTCCTGAGCCACCATCACCTCCTGGGGTGCCAGCACCTG
 TTCTCTTGGTCAGGAGCTGTAGAGATGGAGCTCAAGCACTGGACGACTCTGTCCCCACTGCTG
 GAATAACTCGGGCACAGAGCATGGGACCAAAGTACAGAAAGAGGTTGGGGGAGACCCCCCAG
 CCCTAGACTTCCATCATTCGGGAGACCAACTCAACACCGTCTTTGCCTGAGAACCTGATATATCC
 GTGTTTTTAAATTTTTTTTTTTCTAGCAAAGTTGGGTTTTAATGACTTATGTTTCATAGGAAAC
 CTCTCTGATCCCACACACAAGGAGGGTGATTCTGGGATGAGTTCCTGGTTCTAGGGCATGAGG
 GGCTGGATGGACCTGTCCCCAGGGAGGACATGGCTCTGAGTCCACAGGGCTGAGGAGGCAAT
 GGGAACCTCCCTGGCCCCGGCCCGGTGCTTGTCTCCTCCCCCTCCCACCTCTTCTCCTCCTAGCT
 CCCCAGCTCCCTGCCTATTCCCCCACCTCCGAGGGGCTGCAGCTTGGGAGCCTCCTCAGCAT
 GACAGCTTGGGTCTCCTCCCCAAAAGAGCCTGTCAGGCCTCAAGAACCACCTCCAGGTGGGGA
 GGGCAGTAACGAAAACCATCGCAGGAAATGGCACCCCTCCCTTTTCGGTGATGTTGAAATCATG
 TTAATAATGAAAACGTGCTCTAGGGAAGTGTTCTGTCTCCTCACAGGCTTCACCCACGGCGAT
 GAGGCCCTTGAATGTGGTCACTTTGTGCTGTATGGTTGAGGGACCCTCACACCAAAGGGACCT
 TCCCATGTGAGATGTGCTCCCGCCCCACCTGCCACAAGCAAACACACCACACATGTTCCGGC
 ATGTTGCCCTTTGAACACCCATGAGGACGCCTCCAACCTGCTCTTGGTCTAATAGGGAGTAC
 TGACTGTCAGCAGTGATGATAAAGGAGAGGGGACCCTCTGGTCCCTAGCATGGCACCCAGAGCCT
 CCCCTCTTCTTGTCCTTCAGCCAAAGAGAACTTTCTCTGACTTTGAACTGAATTTAGGTCTC
 TGGCCAATGATGGGCCTGAAAATTCCATAATGGCCAGAGAGGAGAGTTCGAGCCCGGCTAAGA
 TCCCCCTGAGTCATTCTGTGAGGGACCAAGACCCACAGTCCACCAGCCCCAGGGCCCTACCTCC
 TGGAATGCTTTCTTGATCCAGCTTCCCGAAGATCCGACCAGACCCAGGGAGGACGGCACCGC
 TCCGCGGGAGGGAAAGCCAAAGCATGGTGCTTCACCAGCTGGACTCAGGGGCGAGGGGACATG
 GCGCTTGTCAACGTGATGTCTTTTCCACCGTTTCTTCTGTTGATATTCAATGAATC
 CGTCAATCTCTCTGGGAAA

FIGURE 609

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTCTCCCCAGTTCCCCTGTGGGTCTGAGGG
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA
ACAAGTTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCCTGCTGTT
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGT
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
AGCCTCGCCAATTCTTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAG
GAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCAATGATTGTCTTTATGCATCCCC
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT
TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTTACTTGGACATG
AACTTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT
ACAGTAAAAAAAAAAAAACCTTGTAATTTCTAGAAGAGTGGCTAGGGGGGTTATTCATTTGTAT
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAA

615/615

FIGURE 610

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI: 8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNKGATTCATNSHSDSEL
RPEIFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVI
ATNLQEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVF
KNYQTPDHYTLRKISSLANSLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQ
AAVVKALGELDILLQWMEETE

Important features of the protein:

Signal peptide:

Amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites:

Amino acids 192-196;225-229

N-myristoylation sites:

Amino acids 42-48;46-52;136-142

FIGURE 610